ANNUAL CONVENTION TRANSACTIONS

FOURTEENTH ANNUAL CONVENTION

AMERICAN SOCIETY FOR QUALITY CONTROL

ANNUAL CONVENTION TRANSACTIONS 1960

FOURTEENTH ANNUAL CONVENTION

AMERICAN SOCIETY FOR QUALITY CONTROL

MAY 24, 25, 26, 1960

SAN FRANCISCO, CALIFORNIA

PRICE: $5.00

Copyright 1960, By
AMERICAN SOCIETY FOR QUALITY CONTROL, INC.
161 West Wisconsin Avenue
Milwaukee 3, Wisconsin

Litho in the United States of America by
EDWARDS BROTHERS, INC.
Ann Arbor, Michigan

FOURTEENTH ANNUAL CONVENTION COMMITTEE

General Chairman
Harter G. Hudson
Owens-Illinois Glass Company

Program Chairman
Gerald J. Lieberman
Stanford University

Publicity Chairman
Allan M. Hull
United Air Lines

Exhibits Chairman
Frank D. Cardoza
Int'l. Business Machines Corp.

Registration Chairman
Allen Chop
Lockheed Missile Systems

Reception Chairman
Gerald J. Desmond
Gerber Products Company

Special Events Chairman
Walter L. Hurd, Jr.
Lockheed Missile Systems

Arrangements Chairman
Ronald J. Berkheimer
Fibreboard Paper Products

Secretary and Treasurer
Lyle D. Connell
Westinghouse Electric Corp.

1960 ASQC OFFICERS

President
C. E. Fisher
Bell Telephone Laboratories
New York, New York

Vice President
L. S. Eichelberger
A. O. Smith Corporation
Milwaukee, Wisconsin

Vice President
A. V. Feigenbaum
General Electric Company
New York, New York

Vice President
J. Y. McClure
Convair - Division of General
Dynamics Corporation
Fort Worth, Texas

Vice President
Ellis R. Ott
Rutgers - The State University
New Brunswick, New Jersey

Executive Secretary
David S. Chambers
University of Tennessee
Knoxville, Tennessee

Treasurer
Allan M. Hull
United Air Lines
San Francisco, California

Jr. Past President
Leon Bass
General Electric Company
Cincinnati, Ohio

Editorial Board Chairman
Mason E. Wescott
Rutgers - The State University
New Brunswick, New Jersey

FOREWORD

During the early years of the Society, its greatest strength lay in the Sections and Regional organizations, many of which antedated the formation of the national organization. The first four national conventions were held in conjunction with regional conferences (1947-1950).

A major step forward was taken in 1951 when, for the first time, the Annual Convention was held as an independent activity of the Society, completely divorced from any connection with a regional conference. Under the leadership of Wade R. Weaver, the American Society for Quality Control began to emerge as an organization capable of serving the national interests of its members in ways beyond the limitations of section and regional groups. A General Convention Committee was created and given the responsibility to provide the long range guidance and planning needed to insure the programming and management of effective annual conventions. This committee served the Society ably, as the record of nine consecutive conventions under its surveillance amply testify.

With completion of the 1959 Convention, the General Convention Committee was discontinued, thus completing a transition program begun a year earlier to pave the way for the next major step forward in convention management for ASQC, viz, the transfer of the business management of its conventions to the Society headquarters office and the program planning to a committee under the supervision of a vice-president of the Society. Transfer of these assignments has now been completed and the 1960 Convention is the first one to be managed and programmed under this new administrative procedure.

These Transactions constitute a record of the presentations at the Fourteenth Annual Convention. They have not been subjected to review by the Editorial Board of the Society. The timetable for planning our conventions precludes this at the present time. After the Convention, the papers will be reviewed by the Editorial Board and certain of them may appear in Industrial Quality Control.

Request for permission to reprint any portion of these Transactions should be addressed to Professor Mason E. Wescott, Chairman, Editorial Board, The Statistics Center, Blake Hall, Rutgers - The State University, New Brunswick, New Jersey.

While these Transactions are copyrighted, the American Society for Quality Control assumes no responsibility for any of the authors' statements. Responsibility for the content of each paper resides with its author.

Mason E. Wescott
Chairman, Editorial Board

A SHORT SURVEY OF MANAGEMENT'S USE OF OPERATIONS RESEARCH

Thomas E. Caywood, Partner
Caywood-Schiller, Associates, Chicago, Ill.

Operations Research in industry is an accomplished fact. Although it was only a few brief years ago that the concept of applying mathematical techniques to complex business problems and decisions was first pronounced, it has found wide acceptance in the business community. The question then is no longer whether to use Operations Research, but rather where and when.

A large number of case histories pertaining to the more usual variety of Operations Research studies, such as inventory control, production scheduling, transportation problems, equipment utilization and product formulation, among others, have been presented in the journals of several societies interested in the subject. These include the Operations Research Society of America, The Institute of Management Sciences, and Society for the Advancement of Management. Popular business publications such as the Harvard Business Review, Fortune, and a number of weekly business magazines have presented a variety of articles dealing with the achievements of OR.

I would like to take this occasion to review briefly some of those case histories which have not been published elsewhere, and, in particular I have selected cases which are out of the ordinary in order to illustrate the almost endless variety of situations confronting business in which OR has proven itself.

PURCHASING STRATEGY

There are a number of Operations Research studies which have been made which do not fall into the classical areas usually associated with the subject. Such studies frequently are variations of the classical approach either through the definition of objectives or through the combination of two or more classical techniques. In other cases either completely new analytic models and approaches must be developed to correspond to the actual problems under consideration. In these cases, the exacting approach of the Operations Research analyst lends considerable insight into the problem which usually yields worthwhile results.

Some of the non-classical Operations Research studies are obviously qualitative in nature. General studies in the areas of credit screening, purchasing strategies, communication network analysis, and training and development programs fall into this class. In most cases of this type the operations analyst will attempt to develop a quantitative scale along which qualitative comparisons can be made. This quantitative scale may or may not have definite physical or economic meaning.

An interesting study of this variety was completed for a small packing house which deals exclusively with hogs. In comparison to other meat packers, this company is unique because it deals with a tremendous variety of hogs and has been extremely successful economically in comparison to other meat packers. It was decided that Operations Research techniques could provide useful information in the area of purchasing hogs.

In the meat-packing industry, where profit margins are traditionally slim, it is extremely important to understand the relationships among the many factors which affect the profit picture. Equally important is the desirability of knowing the profit situation for short periods of time such as a week or even a day. If one is aware of the basic relationships involved and has the pertinent current numerical information, the opportunities for control are increased.

The meat-packing firm for whom this study was made has considerable flexibility regarding the level of their daily operations. Many plant employees are usually willing to spend an additional hour or two beyond the normal operating

period. The firm processes a variety of livestock with respect to both type and size of the individual carcasses.

Upon examination of the way in which the company operates, it appeared as if the long and short term profit situation hangs completely upon each individual day's activity. This corresponds to the unique situation where the operation is basically on a one-day cycle. While the above phenomenon is not perfect, it is so near perfect that the single day was used as the unit of time in the analysis.

Although the basic decision period consisted of one day, this day was not a calendar day. The decision period consisted of basically the afternoon and morning of two adjacent days. The most critical decision which the company must make each day is the number of each type and each size of livestock to be purchased. This decision hinges upon many factors, and the purpose of this study was to develop a model and a method so that this decision could be made to provide the greatest profit.

It was decided that a set of break even purchase prices would be extremely useful in deciding if specific purchases should be made and also in the evaluation of past purchases. These break even purchase prices have the tremendous advantage that they are easy to use. When using these prices in deciding if a purchase should be made, one merely has to compare the actual purchase price with the corresponding break even purchase price. If the break even price is greater than the actual price, the purchase will result in a profit per hundredweight live equal to the difference between the two prices. If the break even purchase price is less than the actual price, the purchase will result in a loss per hundred equal to the difference between the two prices.

Recognizing the utility of a set of break even purchase prices, the problem consisted of developing a model and method for developing them on a daily basis. The following model was used:

$$p_{ijkl} = y_{ij}(v_{ijk} - c_{ijl}) \quad ;$$

where

p_{ijkl} = the average break even purchase price per cwt. live of the i th size of the j th type for k head of this size and type and l head in total,

y_{ij} = the average live to dressed carcass yield for the i th size of the j th type,

v_{ijk} = the average sales value per cwt. dressed of the i th size of the j th type for k head of this type, and

c_{ijl} = the average processing cost per cwt. dressed of the i th size of the j th type for l head in total.

A description of the analyses which were completed to determine the behavior of each of the three independent variables follows.

The average live to dressed carcass yield for each size and type was developed from historical information and published research studies on this subject. It was observed that there are substantial yield differences among the sizes and types, and this fact supported the decision to consider relatively homogeneous groups of livestock. The yields are obviously independent of the number of head in each group and the total number processed each day.

The determination of the average sales value per cwt. dressed of the i th size of the j th type for k head of this type required a study of the sales

force activities. Here two basic problems had to be considered. First, the average value changes over short periods of time for each size and type with a fixed number of head being processed had to be studied. Second, the average value changes for each size and type with various numbers of head being processed had to be examined. By working closely with the sales department personnel, the answers to both of these questions were obtained. It was observed that short term sales value changes with fixed quantities could be forecasted and that sales value changes corresponding to volume changes could be described with simple rules. A scheme was developed so that on any particular day forecasts of the average value per cwt. for the various sizes, types, and levels of operation could be easily obtained.

The determination of the average processing cost per cwt. dressed of the i th size of the j th type for l head in total required a study of the cost accounting techniques. Fortunately the firm is extremely cost conscious and maintains excellent records regarding the daily costs of the various parts of their operation. They also know how many hours each of the departments have to work for various numbers of head purchased each day. The carcasses are moved through the plant on overhead conveyers which move at fixed rates, and so it is relatively easy to determine the total number of departmental hours required to process various quantities of carcasses.

Since the departmental costs per hour and all other variable and fixed costs were known, it was possible to develop a total cost schedule per day for various levels of operation. Such a cost schedule was prepared on the basis of a processing day, not a calendar day. Since all the carcasses move through the plant at fixed rates, the average cost per carcass is constant and depends only upon the total number of carcasses processed. The average processing cost per cwt. dressed of the i th size of the j th type for each level of operation was obtained by dividing the average processing cost per carcass by the average dressed carcass weight in cwt.

The physical results of the analysis were presented in a series of charts which can be used very easily by the hog buyers. Each chart displays a set of break even prices for the various types of hogs. By using this chart, the hog buyer can determine immediately if the actual purchase price is less than a break even price. If it is the company will realize a profit per hundred weight equal to the difference; and if the actual purchase price exceeds the break even price, he knows a loss will be incurred.

COMMUNICATION NETWORK FACILITIES

The performance evaluation of a teletype system which follows was done for a company in the business of warehousing, transporting, and trading of grain. The conduct of the company's business requires rapid communication between approximately 80 offices in the United States and Canada. This communication is accomplished by a leased-wire teletype system consisting of approximately 20 circuits, most of which terminate in a common relay base. The use of the word "approximately" is not intended to be evasive. The fact that the teletype system is experiencing a phase growth and rapid change due to expansion and changes in the company's business is the very reason for which this analysis was made.

Frequent performance problems arise which generally relate to congestion resulting in delays in message transmission of a degree regarded as undesirable by the management. For each such situation there are usually several alternative actions which may be taken to relieve the problem. An additional circuit may be leased to relieve some of the load on the congested circuit. Existing circuits may be rearranged by leasing different segments of line to redistribute the load among the circuits. In some instances, faster transmission equipment may be installed on a circuit. Also, combinations of these possible alternatives may be considered. With each proposed remedy is associated a penalty in the form of increased line or equipment rentals. Opinions of how a problem situation should be remedied arise at various levels of the management organization and have in the past resulted in conflicts between personalities and solutions based upon organizational rank.

Each circuit is a sort of party line for the stations on the circuit since each station on the circuit may "eavesdrop" on any message transmitted on the circuit. When a message is to be transmitted from one station to a station not on the same circuit, the message must be relayed through a station, usually the relay base, common to the two circuits thus activating two circuits in order to transmit the message. Also messages frequently have multiple destinations. Because of these features of the system, attempts are made to group stations on a circuit on a "community of interest" basis to minimize the relaying of messages. Of course this community of interest grouping must conform with geography, or more specifically, wire lease cost. A message is prepared for transmission by first preparing a perforated paper tape of the message which is then fed into the transmitting apparatus. At a relay point, a message is received in the form of a tape which is fed into a transmitter on the destination circuit. The following few statistics will further describe the system:

5,100 messages per day addressed to 6,800 stations requiring 10,800 transmissions.

Average message length is about 24 inches requiring about 1/2 minute to transmit.

A general model of the entire system would be hopelessly complicated. However, a technique was developed by which some fundamental queuing theory could be applied to an individual circuit to analyze its behavior. In the usual terminology of queuing theory, the circuit constitutes the service channel, the messages to be transmitted on the circuit are the arrival units, and service consists of transmitting a message on the circuit. Once the statistical features of arrivals and service times have been established, the factor which governs the behavior of a service facility, in this case a circuit, is the utilization parameter or fraction of time the facility is busy. In the analysis of the performance of the system, the basic operational problem was the determination of the fraction of time a circuit would be busy. Of course, this was easily determined for existing circuits by direct measurement but was of interest primarily for proposed circuits. The following example will illustrate:

A circuit which linked the stations,

GR Great Falls
KS Seattle
KV Vancouver
RO Portland
SK Spokane
RB Relay Base

was considered with regard to redesign when a few message delays, considered serious, occurred due to congestion of the circuit. The problem was to analyze the following situations:

Is the load on this circuit excessive? If so, what improvement would result from

1) Replacement of the 75-word per minute equipment on the circuit by 100-word per minute equipment.

2) Using present equipment and transferring Vancouver from the circuit to the Winnipeg circuit.

As has been indicated, the operational problem was to determine the utilization parameter for each of the proposed alternatives. This parameter for the existing circuit was obtained by direct measurement using an automatic recording device and was found to be .515 meaning that the circuit was occupied with transmission 51.5% of the time during the operating day. A method, which I shall describe a little later, was devised by which the utilization parameter

could be translated into meaningful operating characteristics. The utilization parameter was established from the statistics shown in Tables I and II, referring to one day's operation.

TABLE 1

Messages Transmitted

	Number of messages
GR	33
KS	32
KV	24
RO	192
SK	34
	315

TABLE 2

Station	Total Rec'd	Duplicates	Reason
RO	110 singles	11	intra from GR
		26	" " KS
		1	" " KV
		28	" " SK
	74 combo	12	combo from GR
		2	" " KS
		5	" " KV
		4	" " SK
GR	16 singles	9	intra from RO
		1	intra from SK
	10 combo	5	combo received also by RO
KS	21 singles	20	intra from RO
	10 combo	6	combo from RO
		3	combo received also by RO
KV	15 singles		
(remove 25)		3	intra from RO
	21 combo	8	combo received also by RO
SK	16 singles	14	intra from RO
	12 combo	7	combo from RO
		4	combo received also by RO
	21 lineups received by all stations		
Total	317	169	

combo = multiple destination
intra = originating on the same circuit

These statistics reveal that

$$315 + 317 - 169 = 463$$

messages were transmitted on the circuit.

Replacement of the 75-word per minute equipment by 100 word per minute equipment would result in a utilization of

$$\frac{75}{100} \times 51.5\% = 38.6\%$$

Removal of KV from the circuit would reduce the transmissions by 24 and net received by

$$15 + 21 - 3 - 8 - 1 - 5 = 19$$

or a total reduction of 43 messages. This would reduce the utilization to

$$\frac{463 - 43}{463} \times 51.5\% = 46.7\%$$

By applying the simplest case of queuing theory, single exponential service channel with Poisson arrivals and perfect infinite queue discipline, it can be determined what percentage of the delayed messages would have to wait, for example, two minutes or longer.

Utilization parameter	% delayed over two minutes
51.5%	9.2%
38.6%	4.2%
46.7%	7.3%

With such characteristics of performance in mind, the subjective reasoning of management was:

The delays in the present circuit are excessive and improvement is required. Installation of 100 word per minute equipment will certainly remedy the situation and allow for future growth. The improvement resulting from transfer of Vancouver to the Winnipeg line is barely sufficient for operating needs and would not justify the cost of a line from Winnipeg to Vancouver.

A more sophisticated model could be developed, at least in theory, by considering the statistics of message arrivals and duration. No feasible means was found for determining the distribution of time between message arrivals on a circuit. However, it was determined that the rate of arrivals remained practically constant throughout the day so that no peak activity considerations were necessary. The distribution of service times was found to be nearly the same for all circuits in the system and to be closely approximated by a type 4 Erlang distribution. A complication in the queue discipline resulted from an informal priority system in which a message might be interrupted to permit transmission of an urgent message. Also an urgent message might be sent by telephone in the event of congestion on the teletype system. A consideration of the above mentioned aspects of the system in the design of the model was not considered commensurate with the subjective nature of the ultimate decision.

MOTOR DESIGN

The rapid growth of Operations Research has corresponded to the growth in the utilization of high-speed computers. Computers have been used in a very great number of Operations Research studies, not as a tool of the trade, but as a tool to save time. Many of the problems we now program in a couple of days and solve in a matter of minutes on the computer could not have been solved by a mathematician or even a team of a dozen mathematicians in time to do any good with the answer. We consider our high-speed, digital computer to be a tool like a slide rule or a desk calculator which is available to our staff to save time.

We do not always use it, but when we do we measure the time we save not in clerical man months but in technical man years.

Recently a study was made of the techniques which engineers use in designing custom built motors, in order to convert these engineering techniques into digital computer language.

Specifically, the problem was to establish the mathematics and devise a program so that when the customer's requirements for a motor were fed into a computer, it would proceed to determine, from the many feasible component arrangements, which was the one best motor design that would satisfy customer specifications at least cost.

There are so many variables and constraints in motor and generator design that it is extremely difficult, when it is not impossible, for a designer to solve directly in terms of the desired performance. A two-speed, two-wind polyphase rotor type of motor, for example, can be designed in several million ways. Thus, in many cases the only feasible way for the designer to proceed is by assuming a certain design and then calculating to see what this design will do in the way of performance. Then he compares the results of these calculations with the customer's requirements and decides what changes might improve the design, and so on until he reaches the final design. The computer performs much the same process until a design which satisfies both performance specifications and economic requirements is reached.

The computer program in the study referred to called for development of the design, including selection of the material and parts of exact or nearest suitability, calculation of required performance data and cost figures, and a printout of total information.

As a result of the program, the design engineer fills out by longhand an input sheet form on which are the customer's requirements for type of motor, rotor core dimensions, rotor dimensions and electrical specifications. A typist enters this information into the computer by typing the hand-written list of numbers. Total elapsed engineer and typist time: usually less than ten minutes.

The computer calls upon the particular routine that is most appropriate, proceeds to compute the optimum machine design, and types out the answer. Total elapsed computer time: less than four minutes on the average. Previously the design engineers in our client company were designing such motors at the rate of one such motor a day.

The study has been of great value in saving valuable engineer time as well as providing answers free of computational error. Since this study, other programs have been established for other purposes. For instance wire size, punching dimensions, and ratings have been used as computer inputs in order to provide flux densities, losses allowed and efficiency achieved, and excitation requirements; or conversely, the computer when instructed with flux densities, losses/efficiency, excitation allowed and starting current, provides wire size, punching dimensions, and weight and cost.

CREDIT POLICIES

With the increase in general credit activity over the last few years, collection from delinquent accounts is becoming more important to many businesses. There are certainly many questions related to a collection operation, and probably most important is the specification of the total amount and type of collection effort which should be allocated to various types of delinquent accounts.

Initially it is necessary to define the collection sequence which usually means enumeration of the basic steps followed in attempting to collect from an account. Each step may consist of sending a form letter to possibly a set of form letters or special dictated letters. Other contact means such as telegrams or phone calls may also be incorporated into the collection sequence.

With the definition of the collection sequence, it is possible to establish the notion of a standard contact. The problem then consists of determining the optimum number of standard contacts to be allocated to various amounts of unpaid balance. The optimum in the case actually consists of the net revenue maximization of the collection activity by developing a balance between the marginal cost of collection and the marginal expected collection revenues.

The analysis under discussion began with two assumptions --

1. the primary objective of a collection activity is to maximize the total amount of profit that is possible from defaulted loans (this is not necessarily the same thing as maximizing returns from defaulted loans), and

2. the department's net profits will be the difference between the amount of money realized from defaulted accounts and the amount of money spent in the attempt to collect on these accounts.

Under these assumptions the most important question then becomes, just how long is it profitable to pursue the non-paying accounts? To pursue such an account too long, i.e., beyond the expected net profit on the account, would be to reduce the department's net profits. Not to pursue such accounts long enough is to reduce, on the one hand, the chance of obtaining the collection, or, where the account is turned over to a collection agency, the total amount of money that might be realized by the company.

The study was made for a large mail order house which sells the majority of its goods through an installment type of buying. The customer opens an account usually with a specific purchase. He is required to furnish certain information concerning his occupation, home ownership status, marital status, number of children, extent of other obligations, etc., upon which he is rated as to how much credit will be extended him. So long as he maintains his account in good order he may make additional purchases to the extent of his total credit rating. The mail order house is interested in maintaining him as a continuing customer. The amount of payment which he will make each month is dependent upon the amount of money owing at the time of each additional purchase.

The first step was to identify delinquent accounts into different types such as skips or non-skips, replies received with partial payment, replies received with non-payment and no replies, first incidence of delinquency, second incidence of delinquency, etc. The analysts then identified the basic collection costs which the mail order house incurred in the various types of contact which a firm collector might use in attempting to collect from the account. These are shown below.

TABLE 3

Means of Contact	Cost in Terms of a Standard Contact
Form Letter	.85
Dictated Letter	1.76
Collect Telegram	.72
Telephone Call and One Form Letter	3.53
Telephone Call and One Dictated Letter	4.45
Two Form Letters	1.12
Three Form Letters	1.40
Collect Telegram and One Form Letter	1.02

An analysis of the historical records in the Collection Department enabled the analysts to determine the patterns of delinquency for each type. For instance for type A (which for purposes of discussion we assume to be a first incidence of delinquency, non-skip, no reply type) the team was able to devise a chart which showed the percentages of the accounts within that type that would remain delinquent as based upon empirical data, during a six-months period.

The analysts found it necessary to identify the profit remaining in an account. In other words, depending upon the amount of money owing the length of time it had been owed (during which time the company made a monthly service charge) and the cost of the efforts made to date to secure payment it could be computed just how much profit remained in the account. Obviously this is an important consideration for justifying the amount of effort to spend. Where the amount of profit is small the amount of effort which can be expended is also small. When the profit approaches or exceeds the hundred dollar mark, a considerable effort is justifiable. From all of the above information a series of tables were constructed. These graphs showed for each type and within each type for each amount of money and each monthly payment area the total number of contacts that the company should make in order to break even.

From these were constructed a number of tables which identify the number of contacts which the firm advises an internal collector to make against certain amount of money owing and according to monthly payment rate agreed upon, for a particular type. After a collector has made the number of contacts listed in the table, if he has not yet secured a payment from a delinquent account, the account is then turned over to an outside collection agency.

The results of the study proved gratifying. The number of accounts assigned to each collector was increased so that the total cost of the department was lowered. Most important, the company was surprised to learn that it had been pursuing small profit delinquent accounts too long and not pursuing high profit delinquent accounts long enough. The final point therefore was not to carry on collection efforts indefinitely (unless for full payment in particular cases), but to be sure to carry it to the point where the amount of profit did not justify further effort by the company itself.

INVENTORY CONTROL

Operations Research analysts have performed a large number of useful studies in determining optimal inventory control policies which have been covered elsewhere; the following case history illustrates a somewhat different kind of inventory control study.

To begin with, it should be understood that the same kind of inventory problem faces the management of a company with large capital items such as aircraft worth from one to ten million dollars each, as faces the management of a company concerned with items such as pieces of paper worth only fractions of a cent each, and that the same rules and controls should apply in both instances.

This case history concerns inventory control in a bank. Unfortunately, or perhaps fortunately for the analysts who made the study, it was not concerned with the control of money, which is another inventory control subject for a bank. Rather, it was concerned with the control of the many and various paper forms used by the bank. It is an interesting case insofar as it reflects the ability of Operations Research to schedule and control the production (or purchase) and storage of diverse items through a single distribution point. In the case at hand the bank had been concerned for a number of years with, principally, 600 forms which were printed on the premises. These forms were used in 144 branch banks in a decentralized organization. Naturally usage of each form varied considerably. Although the bank had many years experience in printing and controlling the storage of these forms, up to 60% of those forms were being ordered through the production shop on a rush priority basis. (Of course, rush priority is costly, especially when delivery must be made to a branch bank 300 miles away.) Costs were substantially increased and interruptions to service were frequent.

The objective of the analysis was the determination and establishment of the policies and procedures necessary to obtain proper control in the central storage point. The importance of a local branch bank control was recognized, but was considered as a supplementary step in the study.

Upon close examination of the forms problem, it was found that control of different types of forms was strongly dependent on variables associated with specific costs of forms as well as probability of usage. As a result, it was possible to group the 600 forms into 16 categories, using 4 categories of costs and 4 of predictability. The cost ratio for each form was then computed. Actual annual usage of a form was used to compute predictability.

One of the most interesting aspects of this case is the obvious simplicity of using the results. The end result of the analysis was presented in a series of sixteen charts, which an inventory clerk can use very easily. The results obtained to date have been significant both in smoothing scheduling operations in the print shop and in providing continuous service to branch bank requests. The whole case constitutes a good example of an instance where it is feasible to apply unit control to very small items.

I hope this brief review of somewhat unusual case histories has served to indicate something of the great number of situations in which Operations Research is serving the business community. The variety of industrial problems to which these techniques have been successfully applied is larger than anyone would have dreamed possible a dozen years ago. And still growing.

WEAPON SYSTEM MANAGEMENT FOR QUALITY

John W. Young
Corporate Director of Quality and Logistics
North American Aviation, Inc.

I appreciate very much the opportunity to talk to you today regarding the role of Quality Control in the management of a weapon system.

As many of you know, our company has had management responsibility experience with three major advanced manned weapon systems and one advanced unmanned weapon system. Now in initial production is the A3J-1 carrier-based attack system for the Navy, which is being produced in our Columbus Division. The F-108 long-range interceptor weapon system was cancelled by the Air Force late in 1959 because of budget limitations, and the B-70 intercontinental strategic weapon system was reoriented by the Air Force early in 1960 to a prototype program for the same reason. Our Los Angeles Division was assigned management responsibilities for these two systems. The GAM-77, an unmanned air-to-surface weapon system now entering the operational phase, is being managed by the Missile Division.

My comments will be principally related to experience gained with the B-70 and the F-108 because of the magnitude of these two programs and also because they represent examples of the current thinking in weapon system management.

I will emphasize the areas where I think Quality Control has increased participation or activity in the weapon system management as compared to our more usual contracts or subcontracts. The areas I have in mind are in the subcontractor selection and contact and also logistics.

A weapon system has been defined as a composite of equipment skills and techniques that forms a self-sufficient instrument of combat. As an example, on such airplanes as the B-70 and F-108 this would include all the components, such as the engine, airframe, electronic equipment, and so forth, and also ground support equipment, maintenance and repair provisions, the training of personnel and all other elements needed to accomplish the assigned mission. The weapon system development concept in turn means adapting and bringing together, in accordance with a plan, various lines of development to achieve such a self-sufficient instrument of combat. The objective is simply to achieve operational availability of the most advanced equipment at the earliest possible date and at minimum cost.

In order to manage the B-70 and F-108 weapon systems, the Los Angeles Division assigned a Weapon System Manager to each program. The Managers reported directly to the General Manager's office and were responsible for all matters pertaining to their particular weapon system. Functional divisions, such as Manufacturing, Engineering, Quality Control, Material, etc., appointed Managers to coordinate and direct the effort expended by the specific functional division on each weapon system. These division Managers, although reporting to their individual division heads, were members of the Weapon System Council and received direction from the overall Weapon System Manager.

The Quality Control function in the Los Angeles Division organized to reflect the importance given to the weapon systems. The B-70 and F-108 Quality Control Chiefs reported to the Director of Quality Control. Each Weapon System Quality Control Chief was responsible for the coordination of quality control activities in his particular weapon system. He was responsible for planning the quality control program, mock-up program participation, and coordinating the weapon system activities in manufacturing development programs and subcontracting activities.

The magnitude of the subcontracting effort in a large weapon system can be appreciated when it is realized that approximately 40% of the dollar volume on the F-108 was airframe, 60% subsystems and GSE, while on the B-70 the percentages were approximately 49% and 51% respectively. Also on the F-108 weapon system, approximately 23% of the dollar volume was to be built by the Los Angeles Division and 77% was to be subcontracted. On the B-70 weapon system, approximately 26% was being built by the Los Angeles Division and 74% was subcontracted.

We can appreciate the job Quality Control must do to help manage such a subcontracting program.

During the initial portion of the Phase I development program of each weapon system, North American prepared contract performance specifications to define the performance requirements of the weapon system air vehicles, subsystem groups, subsystems, and other elements of the weapon system, including training equipment, ground support equipment, ground cooperational equipment, etc. These documents were approved by the Air Force as the contractual documents establishing the performance objectives for the development program. Receipt of initial Air Force approval of these specifications was North American's authorization to solicit proposals from prospective subcontractors for the various subsystems.

To implement the subcontractor procurement program, engineering procurement specifications were prepared. Quality Control, working with Engineering, wrote quality assurance and reliability provisions into these specifications. These provisions incorporated the following basic elements:

1. A clear statement of the subcontractor's responsibility for quality through adequate tests and inspections at his plant.

2. Main categories of tests, such as qualification, quality verification, reliability, and individual acceptance tests, and selection of articles for these tests.

3. Requirement for a time schedule for completion of the major test categories.

4. Methods for reporting test programs and test results.

The review of potential subcontractors for preparation of subsystem bidders' lists was very extensive. It included other prime contractors' experience, the laboratories experience at WADC, and the Los Angeles Division experience. The Purchasing Department solicited recommendations of sources from the Engineering, Quality Control, and Manufacturing departments. The list of tentative bidders was augmented by every available source, including the request of suppliers themselves to have their names added to the list. Efforts were made to assure that all subcontractors currently supplying subsystems in the general category of the requirements were included in the listing.

Prior to starting surveys of potential bidders, the lists were submitted to a Source Selection Board and to the Air Force Weapon System Project Office for review and approval. The Source Selection Board consisted of the Material Director, Chief Engineer, Manufacturing Manager, Quality Control Director, and Weapon System Manager.

Survey teams were then sent into each potential subcontractor's plant to determine his capabilities. The survey teams consisted of qualified representatives of Purchasing, Engineering, Manufacturing, and Quality Control. The Purchasing representative served as team captain. Each team member, and particularly the engineer, was a specialist in the field of the subsystem or subcontract for which the survey was being conducted. A report accumulating replies to hundreds of questions was written at the completion of each survey.

The format of this survey was developed by all departments represented on the team. It was specifically designed to reflect those general and special capabilities of the subcontractor that would serve the need of the system under consideration. In the cost area of the evaluations, it was frequently necessary to conduct a second visitation to the leading contenders' plants to bridge any gaps in cost and price information furnished in the subcontractors' proposals.

After the bidders were chosen, they were invited to a "bidders' conference." The purpose of this conference was three-fold: First, to explain in more detail than was possible in the specifications the significant characteristics of the air vehicle, as well as overall characteristics of the subsystem itself; second, to expand upon those portions of the request for proposal which were peculiar to the weapon system concept of procurement; and third, to emphasize those requirements considered to be of primary importance. There were several by-products of the bidders' conferences which were beneficial to the bidders and ultimately to the weapon system contractor and the government. Among these were: (a) The list of bidders was read aloud at the conference thereby letting each know who was competing; (b) Concepts of weapon system procurement were presented, both in the request for proposal and verbally in detail; and (c) Emphasis was placed on newer requirements under the weapon system philosophy, such as logistics management, ground support responsibilities, data and data processing requirements, etc. A question-and-answer period was scheduled after requirements were presented to permit bidders to present any questions on the material provided by North American Aviation.

Included in the bidder's package was the North American general Quality Control Specification GQ0802-001, which uses the new DOD Quality Control Specification MIL-Q-9858 as its basis but which describes the relationships as they would be established between North American and the supplier instead of between the government and its suppliers.

After the bidders' proposals were submitted, the bidders were invited to make verbal presentations. These were designed to highlight the technical aspects of their proposals and to provide the weapon system contractor a means of verbally asking questions arising from the evaluation. These presentations were scheduled about two weeks after the proposals were submitted to enable the technical evaluating teams time to become familiar with the proposals.

The suppliers' proposals were carefully reviewed by evaluation teams qualified to judge the technical design, technical capability, production ability, probable cost, and program management factors. The weapon systems evaluation teams were composed of Engineering, Manufacturing, Quality Control, and Purchasing personnel.

Engineering team members evaluated the technical design and technical capability factors, and team members from Manufacturing, Purchasing, and Quality Control evaluated the ability to produce, program management, and probable cost. These five major factors were broken down into 164 subordinate factors. The quality control factors considered were primarily under "ability to produce" and "program management." The Quality Control team members evaluated quality control facilities, quality control system, quality assurance program, and the reliability program.

Using a point system of evaluating the proposal as well as data gathered or observed during the earlier surveys, the combined recommendations of the individual teams were presented to the Source Selection Board.

After the evaluations and point ratings were presented to the Source Selection Board, which was the same group that earlier approved the list of those to be invited to bid, the Board made the final selection of the subcontractor. The choice was further reviewed and approved at a formal presentation to the Los Angeles Division General Manager, the corporate management, and then to the Air Force Plant Representative. Final approval was made by the Air Force Weapon System Project Office, which used other Air Force groups, such as

WADC Laboratory specialists, for consultation. The subcontract award was announced by the Los Angeles Division.

During the contract negotiations which followed the subcontractor selection, the Quality Control subcontracting personnel participated in the discussions to make certain there was a clear understanding of the quality specifications and clauses in the contract.

Immediately following the award of the contract, the weapon system contractor assigned specialized personnel to serve as resident representatives at the subsystem contractor's plant. These representatives were given specialized training for the assignment and their role in the total program. One of these representatives was a Quality Control man. His duties were to assist the subcontractor in quality control requirements as defined in the contract by the weapon system contractor's quality control specification or individual quality requirements in numerous specifications. As the weapon system manager, North American Aviation respected the system subcontractor as a full partner in the program; however, the manager must have assurance that quality requirements were being met.

Experience with the assignment of Quality Control representatives to the subsystem contractors' plants was limited by the time the complete weapon systems subcontract programs were in effect. However, sufficient experience was gained to provide confidence in the assigned representative approach to major subcontract control. The representatives were received by the subcontractors with varying degrees of enthusiasm, but after the subcontractors realized they were there to assist and interpret and use surveillance methods they became more cooperative. Also, much of the success of the representatives was probably due to the North American Aviation supplier quality control specification and the effective job the representatives did in explaining it.

There was some feeling that the representatives were assigned to the subcontractors too early in the program. This early assignment soon proved to be a real asset to the subcontractors during their quality control planning activities. Many of the program cost reductions effected by the subcontractors were directly or indirectly the results of Quality Control representative efforts. For example, one subcontractor's practice was to run a complete qualification test on a preproduction system and again later on a production system. The representative was able to suggest redirection of the test program to eliminate the second complete qualification test and substitute a quality verification test program. Renegotiations resulted in a saving of approximately one-and-a-half million dollars.

Another cost saving advantage the representatives provided was in minimizing the data submittal and progress reporting by the subcontractors. The multitude of paperwork that can be unwittingly generated in weapon system management is one of the major and justified criticisms. The representatives had access to information required by North American Aviation and could report much of that considered necessary.

The assurance of quality products other than major subsystems from subcontractors was planned on a basis of surveillance. Quality assurance requirements outlining test procedures, instructions for recording failures and test results, and forwarding of objective quality evidence were included in the procurement specifications. The degree of surveillance to be exercised on a given item of procurement was based on such factors as complexity of the item, its application, the supplier's capability and past performance, testing procedures and results and the frequency and failures during test. Surveillance would range from a minimum inspection in the North American Aviation Receiving Inspection Department to inspection at source.

It must not be forgotten that our customer, in this case the Air Force, has a Quality Control Chief at the weapon system contractor's plant. He has the responsibility to exert surveillance over the effort of the weapon system

managing contractor in the subcontracting program, as well as his effort in the plant. Our Air Force Quality Control Chief informed the Los Angeles Division that the type and extent of his surveillance efforts would depend on the procedures and program the weapon system manager had to assure the quality of the subcontractor's products.

The major problem between the Air Force Quality Control Chief at our plant and the Government Quality Control personnel at the subcontractors' plants was misunderstanding of the weapon system concept and the role the government representatives would have in the administration program. This problem was anticipated early in the program, and North American Aviation and Air Force Quality Control personnel worked as a team to resolve the misunderstandings as soon as they were observed or anticipated.

The Air Force Plant Representative at North American Aviation sent a letter to Air Force Plant Representatives at the major system subcontractors' plants delegating them secondary administration. Teams of Air Force and company Quality Control personnel visited the subcontractor facilities to convey the secondary administration message and set up working relationships. This effort was further augmented by having formal symposia at the weapon system contractor's plant with government representatives from procurement districts, air materiel areas, and the participating subcontractors. These meetings began with an address by Air Force General Officers on the weapon system management concept and the secondary administration approach. After the general meetings were concluded, smaller groups were set up to discuss specific situations and answer related questions.

The Air Force Quality Control personnel at North American Aviation planned to establish regular visits at approximately six-month intervals, or more often if the need was evident, to keep the secondary administration relationship effective.

Also it was apparent that there could be duplication of effort between the Air Force Quality Control and the weapon system Quality Control representatives at the major subcontractors' plants. This duplication was held to a minimum by a close communication and at least weekly meetings with the two parties.

Up to this point, we have considered the role of Quality Control weapon system management in the areas of subcontractor selection, working relations with the subcontractor and Air Force-contractor relationship in the delegation of secondary responsibility. Most of the in-plant Quality Control aspects of weapon system management can be taken in stride, and I will pass over them and into the weapon system operation phase.

All the time and money and work spent on building and delivering a reliable product or weapon system does not pay off until the air vehicle takes off and completes its mission or the missile is launched and destroys the target. The current weapon systems have far greater performance and capabilities and are more complex than our earlier weapons in the hands of the military personnel. These new systems require more ground support equipment, more ground check-out equipment, and these too are more complex. Much of this equipment may be beyond the ability of some military facilities to repair without a probable severe loss of reliability.

Subsystems in the weapon and its supporting equipment must be calibrated, records of operating time must be kept, records of modification must be recorded, work must be inspected, and failure analysis must be made. In other words, there must be quality control.

A simple example of a typical operational problem is the question of compatibility of systems in the weapon with the support and check-out equipment. The weapon system manager has a documentation system whereby systems can be identified as being compatible. This is done by a combination of dash numbers and serial numbers or modification designations. The objective here is to have the operating personnel know what modification of one subsystem will operate

with another subsystem and its modification. Each change to a subsystem is analyzed by the Engineering Department to determine its effect on another, and elaborate tables or charts are prepared and kept up to date to keep the compatibility under control. This effort is for naught if the operating personnel do not keep adequate records and have a quality control system to know the status of modifications.

In most weapon system contractors' plants, the logistics function is charged with customer contact after the weapon is delivered. It is with or through this logistics function that Quality Control has a real responsibility that must be recognized.

The weapon system contractors have for many years integrated a strong and effective quality control function in their operations. With the production of more complex products and systems, the Quality Control personnel have been upgraded technically and received more training. The Air Force Quality Control, mainly in the procurement area, have hired many quality control engineers and assigned them to contractors' plants. I understand by the end of this year they will have well over 300 graduate engineers.

But, is there comparable activity at the operating level to recognize the need for a strong quality control function and plan for its integration?

Here is the field where Quality Control must extend their management knowledge and planning. They should not shut the door and be satisfied that a good, reliable product was delivered.

The contractor's Logistics personnel are well acquainted with the problem the military have trying to maintain high quality and reliability in the systems they receive. I have found the Logistics personnel are very receptive to having their Quality Control personnel team up with them and make suggestions and recommendations to the military. The weapon system manager has the responsibility to impress on the using commands the necessity of training their personnel in quality control and inspection techniques and to impress on them their need for adequate control.

The contractors who have weapon system management responsibilities should assume the lead in this effort. Their subcontractors look to the managers to take on these overall programs.

In addition to being alert to the establishment of effective control in the operating area, Quality Control should assure that important modification kit work is adequately inspected in the field. A planning operation with inspection points similar to those used in the plant should be set up for modification kits. In most cases, a Quality Control man should accompany the modification team sent into the field to make the change.

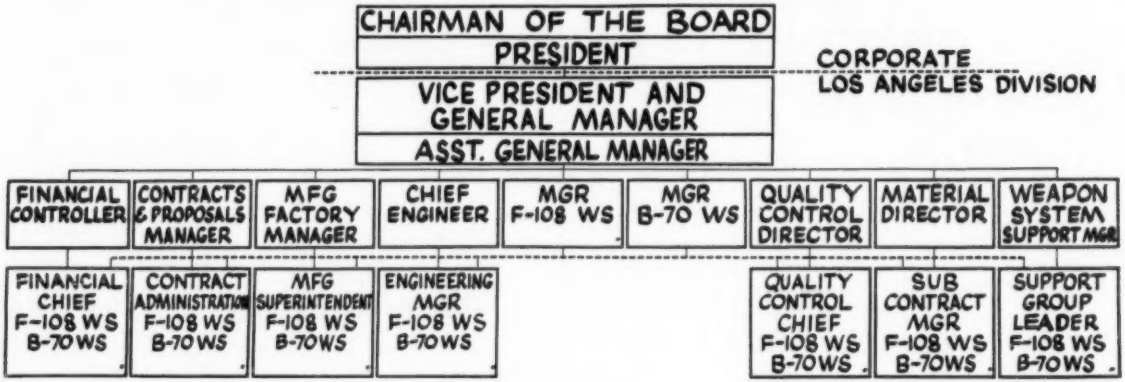
Another practice our company has found effective is to have a Quality Control man make trips to bases operating the company's products to survey and report on the inspection and maintenance practices in use. This has been a great help to Logistics personnel, and the report is also made available to the plant Air Force Quality Control Chief for whatever action he may care to take.

While I have only covered a few of the more important areas in weapon system management where I feel Quality Control has to face up to its responsibility along with the other functions, it is very apparent that there is a big job to be done. To do the job, Quality Control must provide and exercise the technical and management ability the government is buying.

What is needed for the job is good direction and leadership by Quality Control management, backing by top management, and a Quality Control organization of capable and enthusiastic people. This last item usually is the one that needs the most attention. We must, by necessity, compete with Engineering and now with Manufacturing to get more technical people to enter Quality Control as a career. But, let's not forget that schooling, development, and encouragement will make many of our present experienced men very adequate for many of the new assignments.

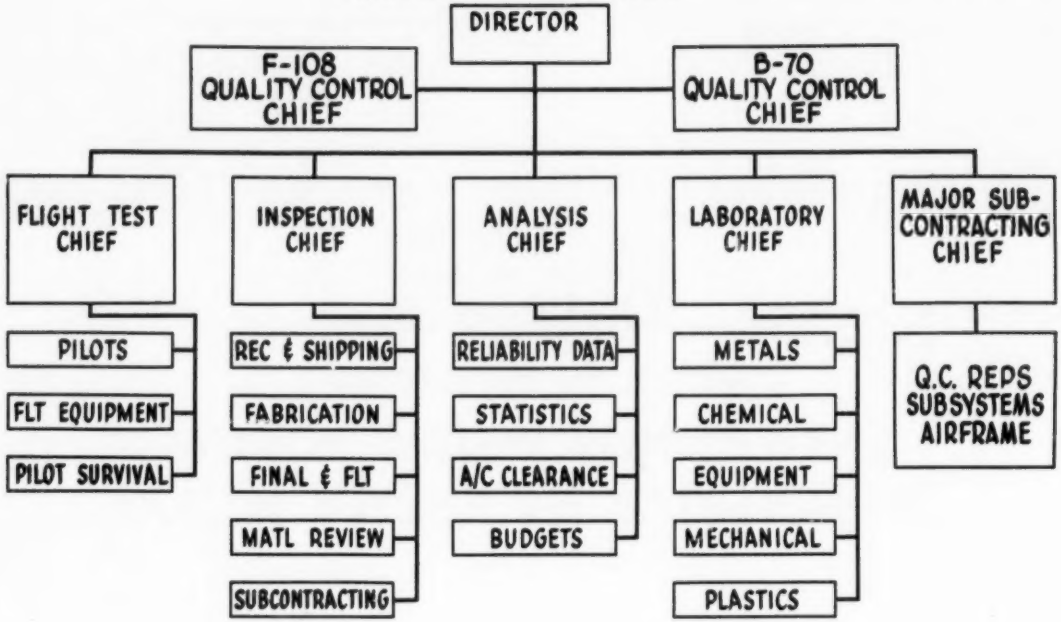
NORTH AMERICAN AVIATION INC.

WEAPON SYSTEM MANAGEMENT


CHAIRMAN OF THE BOARD
PRESIDENT
CORPORATE
LOS ANGELES DIVISION
VICE PRESIDENT AND GENERAL MANAGER
ASST. GENERAL MANAGER
FINANCIAL CONTROLLER
CONTRACTS & PROPOSALS MANAGER
MFG FACTORY MANAGER
CHIEF ENGINEER
MGR F-108 WS
MGR B-70 WS
QUALITY CONTROL DIRECTOR
MATERIAL DIRECTOR
WEAPON SYSTEM SUPPORT MGR.
FINANCIAL CHIEF F-108 WS B-70 WS
CONTRACT ADMINISTRATION F-108 WS B-70 WS
MFG SUPERINTENDENT F-108 WS B-70 WS
ENGINEERING MGR F-108 WS B-70 WS
QUALITY CONTROL CHIEF F-108 WS B-70 WS
SUB CONTRACT MGR F-108 WS B-70 WS
SUPPORT GROUP LEADER F-108 WS B-70 WS

•WS COUNCIL MEMBERS

LOS ANGELES DIVISION
QUALITY CONTROL



THE MEASUREMENT AND SPECIFICATION OF PRODUCT ABILITIES

J. A. Cafaro and H. D. Voegtlen
RCA Service Company
A Division of Radio Corporation of America

INTRODUCTION

Full customer satisfaction with a specific product is not easily achieved. It is dependent on a fairly complex interrelation of the customer's original expectations and the actual performance, dependability, and cost of the product. Reliability, availability, supportability and other product 'abilities are vital characteristics that determine to a large extent the degree of customer satisfaction achieved. The recognition, measurement, and control of these product 'abilities and their interdependence must be considered from the design inception of a product to the end of its operational life. Some of these product 'abilities have been largely neglected in the past. If they have been recognized at all, this has come about only after the product has been placed in the customer's hands, and his dissatisfaction has become known. Quantitative treatment of some important product 'abilities will be considered here, using examples from Air Force electronic equipment studies.

INGREDIENTS OF CUSTOMER SATISFACTION

Certainly we, in the quality control field, are concerned with customer satisfaction. The following definition of the quality control activity, originally stated by A. V. Feigenbaum, I believe, covers the ground:

> "Quality Control is a system for administration and coordination of quality maintenance and quality improvement efforts within the organization so as to enable the manufacture of a product at the most economical level that allows full customer satisfaction."

Figure 1, "Modern Quality Control Defined", presents this definition in a graphic manner and indicates some of the important ingredients in an effective quality control program.

As an idea, quality can be expressed as the degree of conformance to design or customer specifications. Here specifications should include all of the operational parameters, visual and mechanical characteristics, and any other attributes or 'abilities important to the customer. In order to assess this degree of conformance, this total concept of quality, it is certainly necessary to be able to measure quantitatively those characteristics that constitute the overall quality.

The definition above says quite a lot when you read between the lines. For example, consider the most important words, "customer satisfaction." This certainly involves performing in the manner and for the purpose intended. It certainly involves both initial and continued satisfactory performance. It involves conformance to specifications set down by the customer and manufacturer. Unfortunately, it also involves some things that have not, in the past, been specified. For example, consider an electronic equipment: The customer's satisfaction is greatly influenced by the rate of failure of this equipment or any portion of it; it is influenced by the manner or severity of failure when it occurs and its effect on the mission; it is influenced by the time, skills, facilities, and materials necessary to repair the equipment when it does fail; it is influenced by the effort required to prevent failures from occurring; it is influenced by the dollars it will cost to support the equipment for its intended life. We are only beginning to learn how to measure, specify and

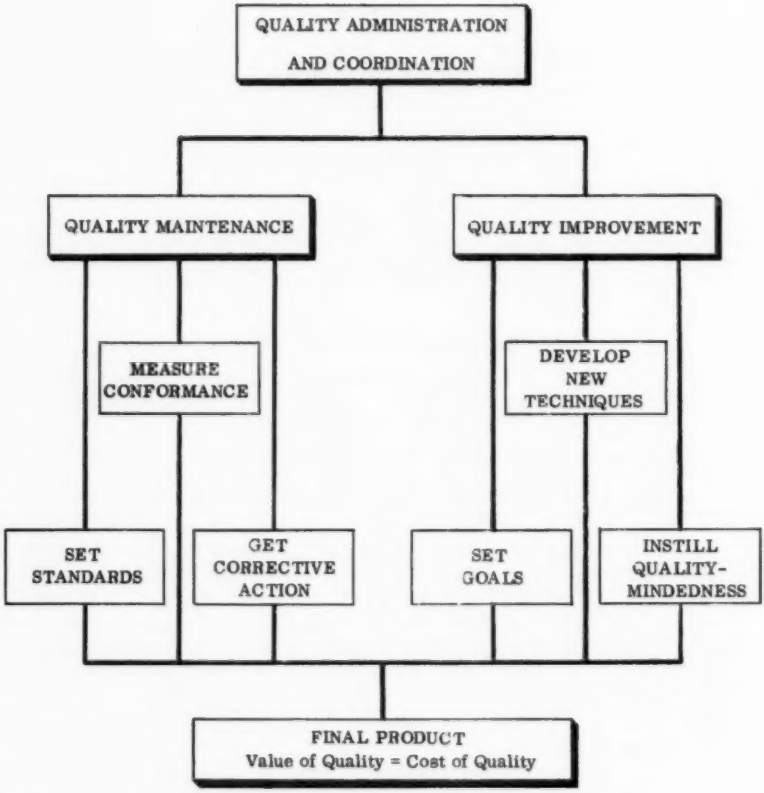


Figure 1. Modern Quality Control Defined.

control these attributes. In fact, some of the above have not even been accurately identified. Nevertheless, they are all quality characteristics of the product and relate directly to the degree of customer satisfaction achieved.

WHAT ARE PRODUCT 'ABILITIES?

Some important product 'abilities are reliability, repairability, supportability, serviceability, operability, produceability, etc. Do we normally measure these product characteristics? Are they stated quantitatively in specifications with compliance demonstrations called out? Are we even able to identify the principal factors that make up these 'ability terms? Except for reliability, which is now recognized as a prime numerical system requirement in many equipment specifications, the answer is "no". These characteristics are generally recognized in a qualitative sense, but little attention has been paid to expressing them in numerical quantities. This was true also of the term, reliability, only a few short years ago. We now have some reasonably well defined terms such as failure rate, mean time to failure, and probability of survival that can be expressed quantitatively and for which compliance demonstration is frequently possible. But how shall we measure such attributes as maintenance skill levels, equipment accessibility for repair, and adequacy of test equipment or instructions, all important ingredients of repairability?

RELATING PRODUCT 'ABILITIES

It may be helpful at this point to consider the overall requirement of a product. A customer's satisfaction depends on how well the product does its intended job, its net worth or value. We may express these qualitative items under the general term "product effectiveness". Effectiveness (E) is a function of performance (P), dependability (D), and cost (C).

$$E = f(P, D, C) \qquad (1)$$

Performance will include the capacity to meet equipment parameters such as range, power output, sensitivity, accuracy, etc. Dependability is the degree of consistence of performance and will include such terms as reliability, maintainability, availability. Cost must include the total amount for development, production, and service life support. In every equipment, some balance among these three factors must be achieved, based on cost considerations, mission requirements, allowable time to develop, etc. For example, unusual performance requirements well beyond existing state-of-the-art technology will result in longer development time and higher cost. Unusual requirements for availability may necessitate redundant components and increase initial cost and support requirements. The process of weighing required performance against the economic factors involved must be carefully considered very early in the development program.

In relating system 'abilities to product effectiveness, it may be useful to draw important inferences from several illustrations. In Figure 2, system or product effectiveness is depicted by a rather decrepit frame building. This is symbolic of the status of many systems which are designed and produced with little regard or control of product 'abilities. In our illustration, product 'abilities are shown as the basic foundation blocks that support the effectiveness of a system and its requirements for operational support during its intended life. Many of our electronic equipments designed within the past decade give evidence of weak foundations. Typical of this, is our foundation block labeled 'Design for Reliability". Often the lack of design goals or misinterpretation by either the customer or the contractor results in a poor foundation in this respect. For example, studies of representative military electronic equipments in wide usage today reveal the following facts:

> For every 1000 hours of operation, three primary equipment failures occur for each 1000 parts of equipment complexity. At this rate, a complex non-redundant system of 35,000 parts (not unusual today) will experience a primary failure every ten hours of operation.[1]

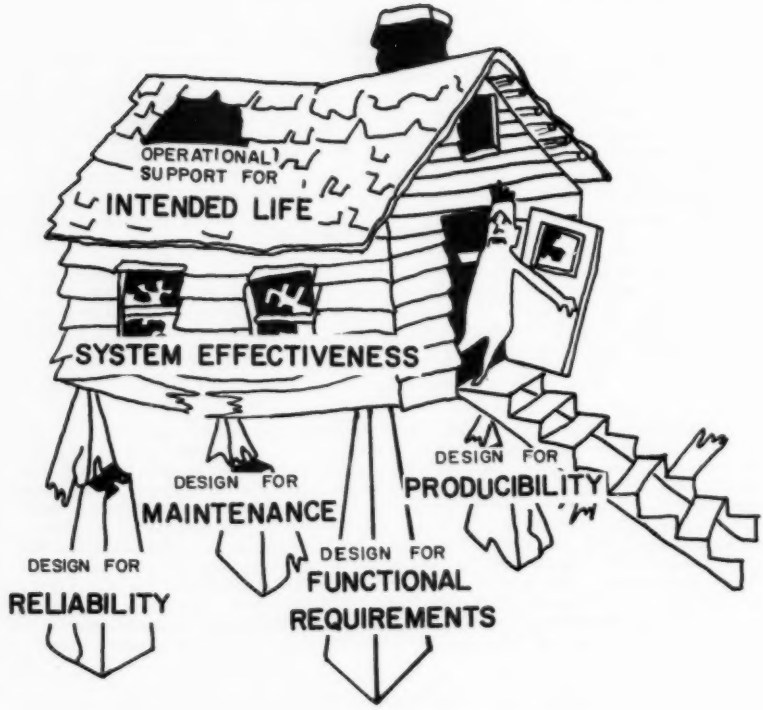

OPERATIONAL SUPPORT FOR
INTENDED LIFE
SYSTEM EFFECTIVENESS
DESIGN FOR
RELIABILITY
DESIGN FOR
MAINTENANCE
DESIGN FOR
FUNCTIONAL
REQUIREMENTS
DESIGN FOR
PRODUCIBILITY

Figure 2

> And this is a conservative figure, representing ground equipments in mild environments!

Consider now the foundation block labeled "Design for Maintenance". Where definite steps have recently been taken to specify reliability in quantitative terms, little similar work has been pursued in the area of specifying design requirements for maintenance (serviceability, repairability). This has resulted in many present day systems which experience excessive down time and high overall support costs in terms of manpower, material, and test facilities. Recent studies of some representative military electronic equipments indicated the following:

> The maintenance man-hour requirements on typical ground equipments amounts to 30 man-hours for every 1000 operating hours for each 1000 parts! This is equivalent to one man-hour maintenance for every hour of operation on the previously mentioned non-redundant system of 35,000 parts.[1]

Another foundation block supporting our system effectiveness structure is "Design for Producibility". Many large electronic systems may be "one of a kind", and therefore, pose a different problem from the production of large quantities of less complex equipments. However, in both instances there must be ample consideration by design of how the production facility can carry through the design objectives. For example, mechanical arrangements which require excessive tool-up time should be eliminated or modified in keeping with design objectives. The design should be thoroughly reviewed prior to release to the production department, so that those areas contributing to needless production cost can be eliminated.

Referring again to Figure 2, we notice that the foundation labeled "Design for Functional Requirements" is well laid. This is symbolic of the highly definitized and quantitatively expressed specifications concerning range, power output, sensitivity, and other parameters predicated on overall mission requirements.

You will notice in Figure 2 that a rather poor situation exists with regard to the roof of our structure. The roof represents "Operational Support for Intended Life", and ideally should protect the integrity of our structure (or system) in its field environment. It is symbolic that our roof in this case is in dire need of repair, as, indeed, are many of our support programs which are inadequate or economically unsound. Recent studies have shown that:

> It costs one Service $40,000 each year to support each 1000 parts of equipment complexity! This figure includes only maintenance manpower and material; no test equipment, depots, housing, or other fixed installations are included.[2]

Admittedly, much of this high cost of support is due to unrealistic equipment design. Economic protection of design integrity requires a support philosophy conceived early in the design phase. During this period it is possible to establish design trade-offs which provide for more realistic support programs.

Figure 3 shows a mature system where the essential product 'abilities have been treated quantitatively and controlled from the inception of design up to the point of service use. Symbolically, this is illustrated by the solid foundation and structure. Unfortunately, the illustration also shows that, even though we have controlled product 'abilities and produced an effective system, inadequate support negates the ultimate goal. Use of the roof as symbolic of "Support for Operational Life", reflects on the fact that no system, however well designed and produced, can be expected to adequately fulfill its mission capabilities unless protected (so to speak) in its field environment by a carefully planned and executed support program. For many of our present equipments we have failed to provide suitably trained maintenance personnel, maintenance instructions are hopelessly confusing, spares are inadequate just

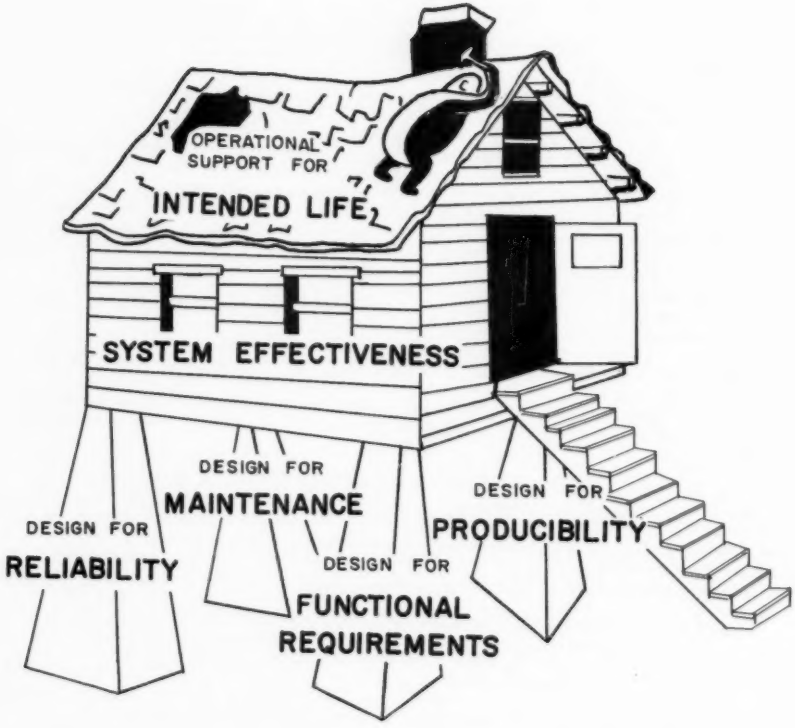

OPERATIONAL
SUPPORT FOR
INTENDED LIFE
SYSTEM EFFECTIVENESS
DESIGN FOR
RELIABILITY
DESIGN FOR
MAINTENANCE
DESIGN FOR
FUNCTIONAL
REQUIREMENTS
DESIGN FOR
PRODUCIBILITY

Figure 3

where the critical need exists, and test equipment is missing or out of calibration. In still other instances changing requirements have resulted in use of existing equipments under conditions for which they were never intended.

Figure 4 is in many respects similar to the previous illustration. However, a well planned support program now protects the integrity of our design in its field environment. Admittedly, this is an optimistic picture in these days of rapid technological change. However, real advances can only come when we are able to identify, measure, and control these important product 'abilities, relate them to other prime product requirements, and plan and implement an adequate support program.

EXAMPLES - MEASURING PRODUCT 'ABILITIES

The following examples illustrate quantitative treatment of three product 'abilities using typical Air Force equipments in the communication, navigation and radar classes. The equipments are prime to the Aircraft Control and Warning mission of the Air Defense Command. They are required to be operational 24 hours a day. But before looking at the data, let's expand equation (1) to gain some perspective. Figure 5, "Ingredients of Product Effectiveness", isolates a few of the major product 'abilities and some of their relevant factors.

Reliability

The observed reliability of an electronic equipment is affected by a number of factors. The principal factor of this figure of merit is the inherent capability of the design; however, usage, maintenance and other environmental factors may have a marked effect on measured reliability. A valid and useful field measurement of reliability must list and measure the factors believed to have a significant effect on the results. A useful standard has been prepared by the Electronics Industries Association[3] for reporting reliability measurements. This guide was followed rigorously in the collection and analysis of data presented here.

In summary, the measure of reliability used is mean time to failure. It is defined as the average time between random failures which cause the effective loss, or degradation below specification limits, of at least one of the essential functions of the equipment. Thus, secondary or non-operational failures, preventive maintenance replacements, etc. are not included in the reliability measurement.

Table I, "Summary of Field Reliability Measurements", presents the information for a sample of four equipments of each class. Confidence intervals for observed failures were derived from the Poisson Distribution at the 90% confidence level. This was the basis for calculation of upper and lower values for the mean time to failure estimate. Use of this distribution assumes that the occurrence of failures are distributed as a Poisson, i.e., the failure rate will be relatively constant over the period of observation. The validity of this assumption was tested for each equipment by application of suitable tests for goodness-of-fit of the observed data to the exponential pattern. Such a test is shown in Figure 6, "Reliability Function (AN/GRC-27, Site 4)".

OPERATIONAL SUPPORT FOR
INTENDED LIFE
SYSTEM EFFECTIVENESS
DESIGN FOR
RELIABILITY
DESIGN FOR
MAINTENANCE
DESIGN FOR
PRODUCIBILITY
FUNCTIONAL
REQUIREMENTS

Figure 4

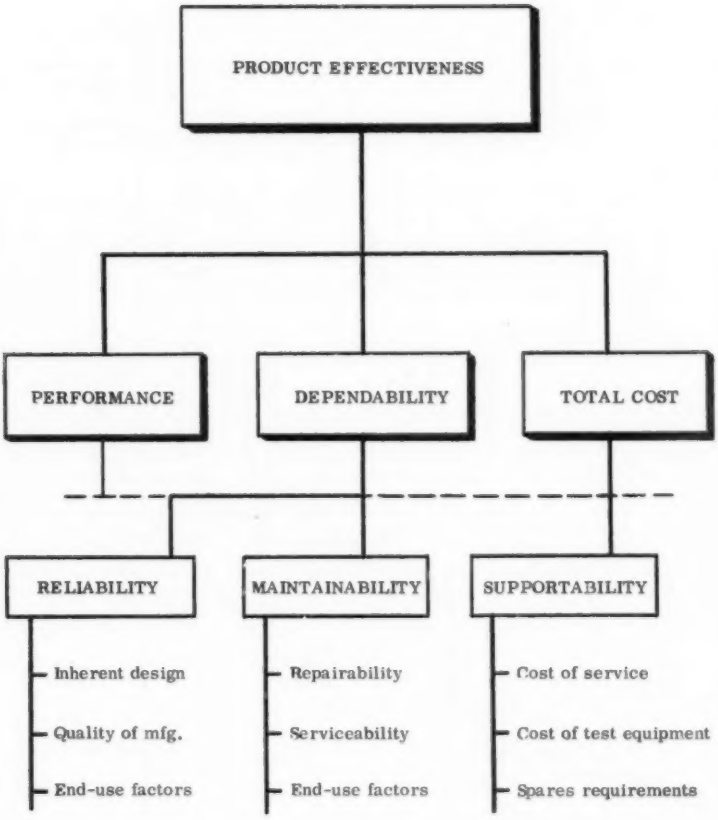


Figure 5. Ingredients of Product Effectiveness.

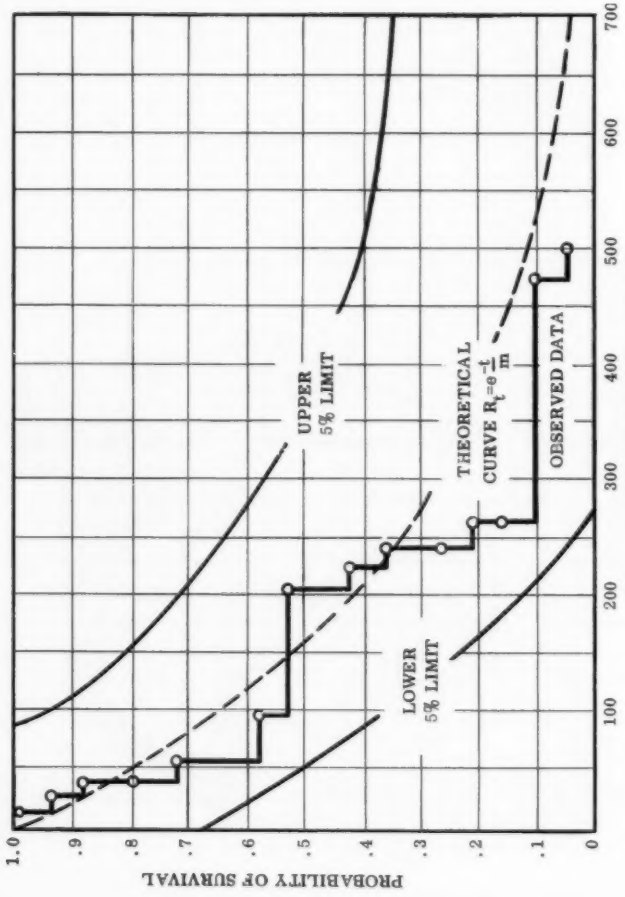


Figure 6. Reliability Function (AN/GRC-27 Site 4).

Table I - Summary of Field Reliability Measurements

	Radar	Communication	Navigation
Complexity (Parts)	6100	1600	540
Total Operating Time (Hrs.)	20960	21076	20887
True Random Failures	382	54	62
Mean Time to Failure Estimate (Hrs.)	55	390	337
90% Confidence Intervals (Hrs.)	51-60	319-490	278-418

From these results an overall Reliability Index (R.I.) has been derived for the three equipment types. This index is 285 bits (approximately .3% failures/1000 hours), and is derived as follows:

$$R.I. = \frac{\text{Total failures}}{\text{Total parts x operating hours}}$$

Figure 7, "Reliability Index - Air Force Ground Electronic Equipment", presents this information on a convenient frame of reference frequently used for reliability assessment.

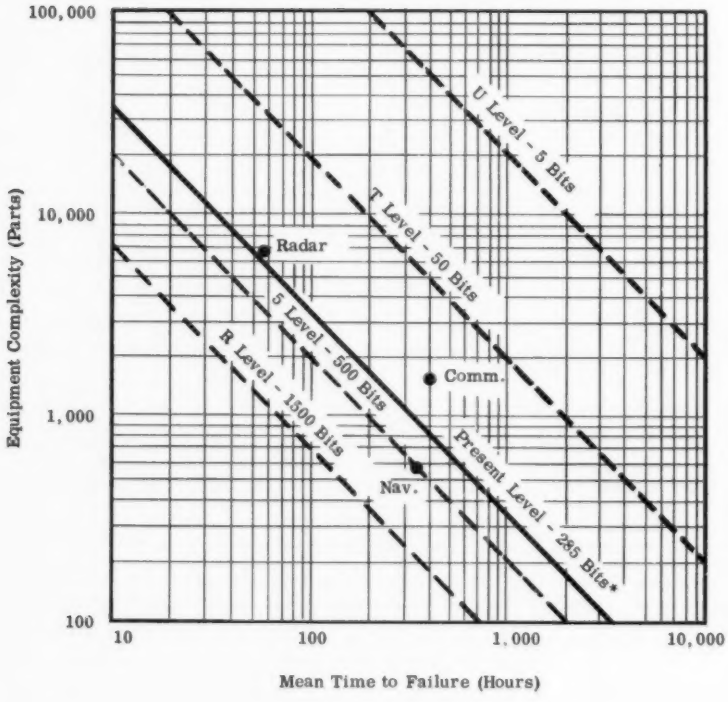
Maintainability

Table II, "Elements of Maintenance", summarizes maintenance information for the three equipments. Referring to this table, down time includes both scheduled and non-scheduled maintenance, waiting time, etc. The logic of associating down time with part replacements derives from the fact that over 95% of all scheduled and non-scheduled maintenance periods involved the replacement of one or more parts or units. Down time per replacement is converted to man-hours by considering the average number of maintenance men directly involved in a maintenance action. For the radar an average of four men were used, for the navigation and communication sets an average of two men were used during maintenance.

Table II - Elements of Maintenance

	Radar	Communication	Navigation
Operate Time (Hrs.)	20960	21076	20887
Down Time (Hrs.)	925	561	231
Parts Replaced	1256	185	129
Repl./1000 Hrs.	60.0	8.8	6.2
Average Down Time/Replacement	.74	3.0	1.8
Average Man-Hours/ Replacement	3.0	6.0	3.6

Figure 8, "Maintainability Index - Air Force Ground Electronic Equipment", provides a useful frame of reference for expressing the degree of maintainability achievement. It has been derived as follows:



* Bits = 1/1000 x % failures/1000 hours

Figure 7. Reliability Index - Air Force Ground Electronic Equipment.

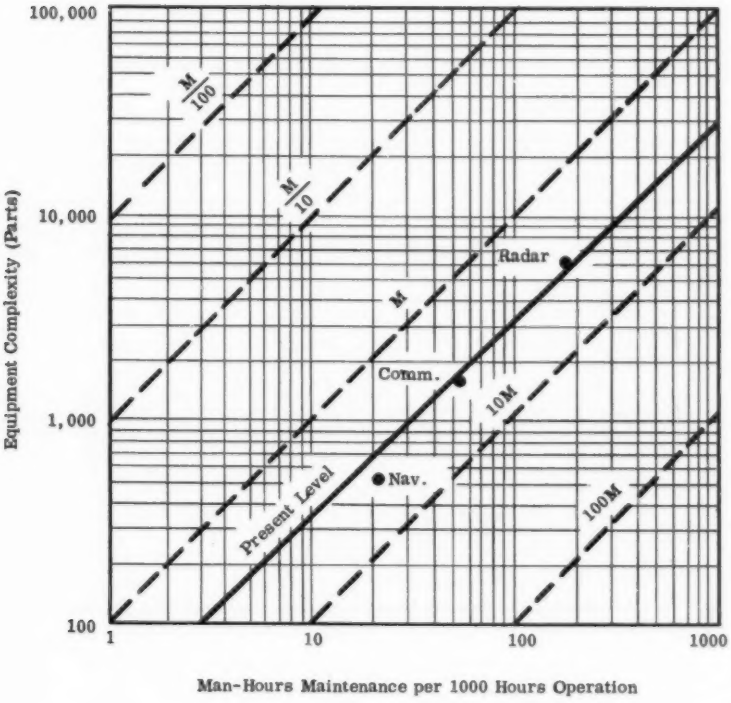


Figure 8. Maintainability Index - Air Force Ground Electronic Equipment.

Step 1 - From Table II, replacements per 1000 hours of operation for the radar is shown to be 60.

Step 2 - From Table II, average man-hours per replacement for the radar is shown to be 3.

Step 3 - Average man-hours per 1000 operating hours is, therefore, 3 x 60 or 180 hours.

Step 4 - Referring to Figure 8, the radar is plotted opposite 180 hours on the horizontal scale and opposite its average complexity in field use (6100 parts) on the vertical scale.

Step 5 - Similarly the other two equipments are plotted and a line of best fit is drawn.

The solid line labeled, "Present Level", represents the current level of maintainability for Air Force ground electronic equipment of the three classes studied. It is quickly noticed that there is a linear relationship between units of maintenance (i.e., man-hours) and equipment complexity. Even though the equipments differ by an order of magnitude in complexity and are completely different in function, the maintenance rate is nearly the same, namely, three man-hours maintenance per thousand hours per 100 parts of equipment complexity. This is, therefore, an index of maintainability appropriate to the equipments studied. Various maintainability levels are plotted on Figure 8 in orders of magnitude (10X). They are intended to provide a frame of reference on which other equipments can be plotted and compared.

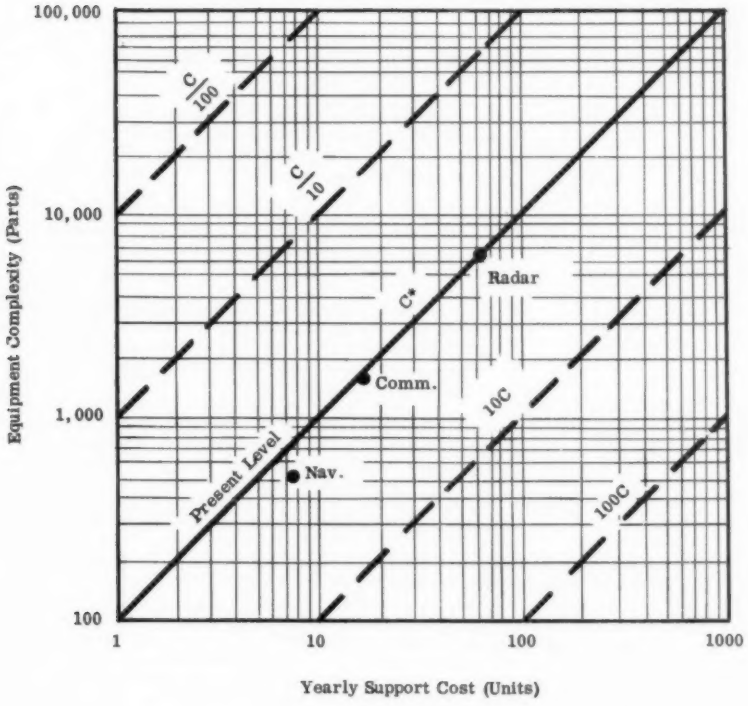
Supportability

Some estimates of support cost for the three equipments have been derived.[2] These estimates include only direct maintenance manpower and material costs. Fixed installations, storage facilities, test equipment, etc. are not included. Table III summarizes this information in cost units, with a comparison of original equipment purchase price. (One cost unit = $4,000.00).

Table III - Equipment Procurement Cost Vs. Yearly Support Cost (Units)

	Equipment Cost	Maint. Manpower Cost	+ Material Cost	= Yearly Support Cost	Yearly Dollar Cost for Support
Radar	100.0	57.0	3.0	60.0	$240,000
Comm.	1.5	17.2	0.3	17.5	70,000
Nav.	1.2	7.2	0.16	7.3	29,000

Figure 9, "Support Cost Index - Air Force Ground Electronic Equipment", provides a convenient frame of reference for expressing support cost levels. It has been derived by plotting the support cost units for each equipment (Table III) opposite the respective equipment complexities. It can be readily observed that there is a linear relationship between support cost and complexity. At the "Present Level" shown in Figure 9, this yearly cost of support is equal to one cost unit for every 100 parts of equipment complexity. Various cost levels are plotted in Figure 9 in orders of magnitude (10X). They are intended to provide a reference for comparison of other equipments in the ground, shipboard or airborne class, and as a guide in evaluating new proposed designs prior to development.



* One Cost Unit = $4000 (Yearly cost of support/100 parts of complexity)

Figure 9. Support Cost Index - Air Force Ground Electronic Equipment.

THE TASKS AHEAD

The measures of maintainability and supportability described here are crude indeed. Both items contain numerous individual factors that make up the end values presented. These factors in turn must be identified, measured, controlled and improved. This orderly process of effort will be required for each of the major product 'abilities.

As an example, we may outline the following steps that constitute the total effort required for each major product 'ability:

(1) Identification - What are the principal factors that limit product effectiveness. These factors must first be identified.

(2) Measurement - How can these factors be expressed in numerical terms? This is necessary for an understanding of their quantitative effects on product effectiveness.

(3) Specification - If a factor (or product 'ability) is important to the user of a product, it must be placed in specifications, with suitable means for compliance demonstration.

(4) Control - When product 'abilities are recognized, measured and incorporated in specifications, a reasonable basis for control can be established, extending from product inception through development, production, and field use.

(5) Improvement - Only after achieving a degree of control of the quantitative levels and variability of these factors will it be practical to improve them. This is the end objective and should be accomplished in the order of priority of their relative contribution toward product effectiveness.

Figure 10, "Recognition and Numerical Treatment of Product 'Abilities", illustrates this cycle of effort required.

We must recognize that product 'abilities, though shown in the three examples as end-of-the-line measurements, are really fundamental building blocks. They come into being when the product's mission or function is first determined, they are shaped and colored by the design concepts embodied in development, and are directly affected by all the factors of end-use environment. Figure 11, "Steps in Logical Design for Product 'Abilities", illustrates this cradle-to-grave concept. It is recognized that, at the present state-of-the-art, there are no universally accepted measurement techniques for some product characteristics such as maintainability and supportability. It is apparent that many comprehensive studies and the extensive collection and analysis of data will be required before generally accepted criteria will prevail. These efforts require a coordinated approach across organizational lines; they involve interdisciplinary actions among many specialists. Many of our existing organizational structures within industry and the military are not presently geared to allow mature treatment of product 'abilities.

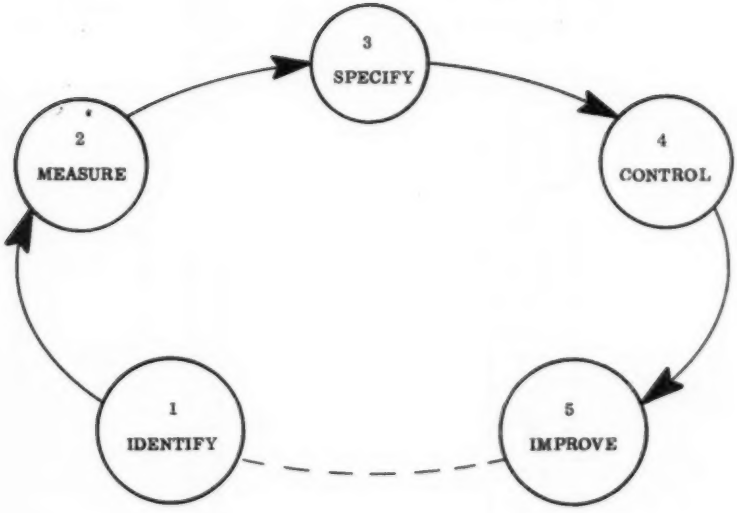


Figure 10. Recognition and Numerical Treatment of Product 'Abilities.

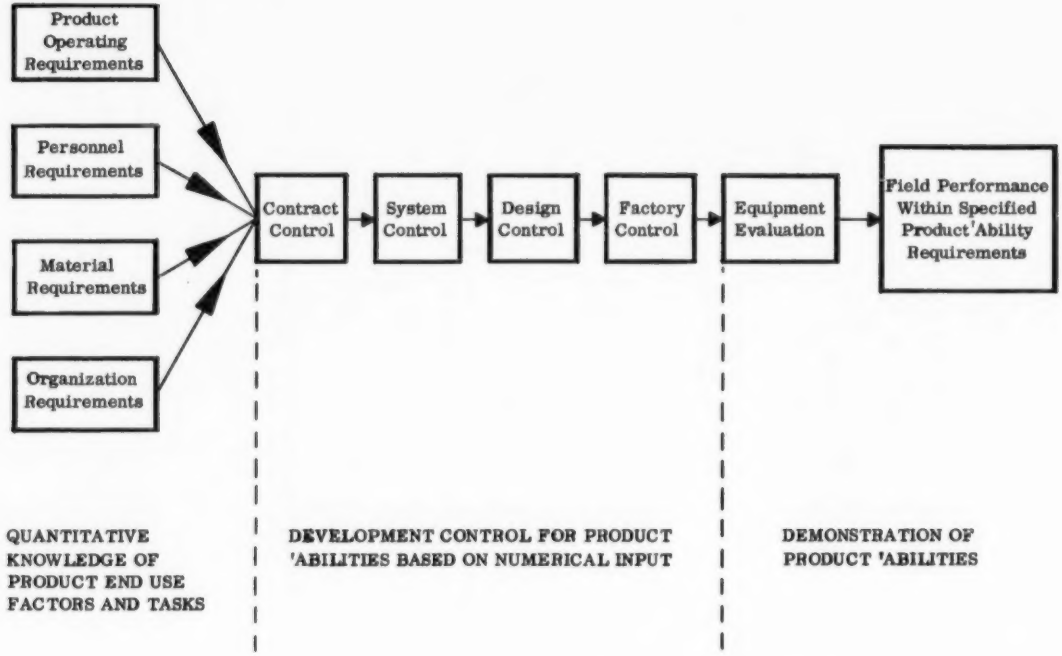


Figure 11. Steps in Logical Design For Product 'Abilities.

REFERENCES

1. "The Prediction and Measurement of Air Force Ground Electronic Equipment Reliability", RCA Service Company, A Division of Radio Corporation of America, sponsored by Rome Air Development Center, ASTIA #AD-148977, August 1958.

2. R. L. McLaughlin, H. D. Voegtlen, "Ground Electronic Equipment Support Cost Vs. Reliability and Maintainability", Fifth National Symposium on Reliability and Quality Control, Philadelphia, January 1959.

3. "A General Guide for Technical Reporting of Electronic Systems Reliability Measurement", Electronic Industries Association, M-5 Systems Committee Report, December 1956.

WHEN IS A SMALL SAMPLE LARGE ENOUGH?

Edward P. Coleman, Professor of Engineering
University of California, Los Angeles

INTRODUCTION

Man in his history has acquired knowledge by logical reasoning and by experiment. Feller in reference [3] of the list at the end of this paper delineates the essential properties of these two processes very well indeed. In the first, he refers to the establishment of new facts or knowledge by operating on accepted notions with precise, unchallenged rules. The traditional subjects of geometry, algebra and later probability, with their particular definitions and axioms are illustrations of this process of learning by logical reasoning and it is independent of any relation to reality. In the second of these processes, Feller emphasizes the importance of experiment or observation. The manner in which experiments are carried out does not depend on preconceived mathematical theories, but it is a purposeful technique depending on, and changing with experience. The history of the development of knowledge shows a stimulating interplay between these two processes, i.e., theory and application. In this interplay, intuition has played a strong role in advancing knowledge. The collective intuition of mankind appears to change with experience and to progress with time.

In almost all investigations involving systems containing human beings and elements of the natural world, it is rare indeed if complete knowledge of the system is known or can ever be determined. However, in spite of this state of incomplete information, today's investigator must make decisions in the face of this kind of uncertainty. Decision making in the face of uncertainty is in part a statistical technique, and its applicability ranges from almost all inductive sciences to many situations that people face in everyday life when it is not perfectly obvious what they should do. Executives, scientists and engineers working in modern industry cannot escape the problem of making decisions in the face of uncertainty.

Let us consider the following catagories of uncertainty and ask certain questions:

1. Uncertainty Due to Basic Design Changes and to Randomness:

   From the results of testing under required conditions pilot* equipment undergoing basic design changes, can one estimate with a high degree of confidence or make statements with low risks of wrong decisions about a performance parameter of the ultimate equipment?

2. Uncertainty Due to Many Small Engineering Changes and to Randomness:

   From the results of testing under required conditions prototype** equipment undergoing many small engineering changes, can one estimate with a high degree of confidence or make statements with low risks of wrong decisions about a performance parameter of the ultimate equipment?

3. Uncertainty Due to Known Differences in Design and to Randomness:

   From the results of testing under required conditions several prototype equipments built to determine the most feasible design configuration, can one estimate with a high degree of confidence or make statements with low risks of wrong decisions about a performance parameter of the equipment?

* Pilot refers to a trial unit of equipment used in testing preparatory to fixing the design and full scale production.

** Prototype refers to the earliest unit of equipment that may be classed as designed to full specifications.

4. Uncertainty Due to Restricting the Test Environment and to Randomness:

   From the results of testing specified equipment in an environment restricted to achieve economy in time, money and manpower, can one estimate with a high degree of confidence or make statements with low risks of wrong decisions about a performance parameter of this equipment operating in the required environment?

5. Uncertainty Due to Modifying the Test Procedures and to Randomness:

   From the results of tests modified to accelerate severity or to substitute non-destructive for standard destructive tests on specified equipment, can one estimate with a high degree of confidence or make statements with low risks of wrong decisions about a performance parameter of this equipment operating under the required conditions?

6. Uncertainty Due to the Derating of Components and to Randomness:

   From the results of testing equipment using some derated components in an effort to meet long life requirements, can one estimate with a high degree of confidence or make decisions with low risks of wrong decisions about a performance parameter of this equipment using derated parts?

7. Uncertainty Due to Faulty Test Instruments and to Randomness:

   From the results of testing specified equipment with faculty test instruments, can one estimate with a high degree of confidence or make statements with low risks of wrong decisions about an actual performance parameter of this equipment?

8. Uncertainty Due to Randomness:

   From the results of testing with proper instruments under required conditions a given number of units of specified equipment selected randomly from the entire supply under study, can one estimate with a high degree of confidence or make statements with low risks of wrong decisions about a performance parameter of the entire supply of equipment under study?

These questions identify the most important sources of uncertainty which affect decision making. While they are stated here in somewhat statistical language, you will hear them referred to in industrial organizations by many vague and cryptic remarks such as, "This is only an experiment model", "Too many engineering changes", "The design has not settled down", "We are still experimenting with the design, yet", "The test environment is not representative", "The tests used are not real", "The contract would be over before we could actually test", "The tests would cost more than the equipment itself", "The test instruments are no good", "How can we get information from one", "We do not have enough to sample", etc.

It is little wonder that individuals even with the best of minds tend to be confused about these sources of uncertainty in making decisions about modern complex equipment, if they have not given serious thought to sound engineering and modern statistical techniques. Randomness is an intrinsic contributor to uncertainty in all eight questions. There are seven sources of uncertainty other than randomness that may or may not be present. However, these sources are not all mutually exclusive, and they are not necessarily additive in cumulative effect. There could be as many as 127 mixed or combined sources of uncertainty. It is apparent therefore that no matter how few sources of uncertainty may be acting on a complex system in test, they are not readily identifiable from the test results themselves. Sound engineering analysis and practice can prevent or reduce the effect of some of these sources of uncertainty. Properly designed statistical experiments can help to estimate what to expect from randomness and to make decisions in the face of uncertainty about other contributing factors or causes of variation. Excellent statistical techniques oriented to the design of experiments point of view are shown in [1].

EXPERIMENTAL BASIS FOR DECISION IN THE FACE OF UNCERTAINTY

Mood in [5] has appropriately stated that statistical inference is based upon two principles of the scientific method: (1) the performance of experiments and (2) the drawing of objective conclusions from experiments. The word experiment is used in many contexts. Therefore, let us describe several types of experiments to emphasize the important connotations in industrial experimentation:

Random Experiment

In a random experiment, emphasis is placed upon the random method of selecting from the entire population or supply under study, an individual unit, object, item or element. A random experiment can be thought of as a real, idealized or conceptual procedure of selecting a unit in which each unit of the entire population is given an equal chance or opportunity to be selected. The actual random selection of a unit can be made by placing all units under study in a bowl, stirring them thoroughly, and then drawing out a unit while one is blindfolded. This procedure of selection can be simulated by use of a random number table and by other mechanical schemes. A random experiment can be repeated for a finite or infinite number of times by repeating the prescribed selection procedure under the same physical conditions. Each such repetition or trial of the random experiment will result in the selection of a different or the same unit from the whole group or population.

Engineering Experiment

In an engineering experiment, emphasis is placed upon the engineering methods of testing and evaluating a selected unit, object, item or element. An engineering experiment can be thought of as a predetermined procedure for testing, observing, measuring or assigning a quantitative value to one or more variables, quantities or quality characteristics of the selected unit. The type and number of quantities to be measured or evaluated for a unit under test should be designated only after consideration has been given to cost and the usefulness of the results. An engineering experiment can range from the visual inspection of a simple unit of equipment requiring only a fraction of a second to the complicated application of expensive measuring instruments making hundreds of observations on a complex system and possibly requiring years to complete. Each quantitative value is recorded to a fixed degree of accuracy and is usually not repeated on a single unit. If the test instruments are in a proper state of calibration, repetition of the engineering experiment on a particular unit would merely duplicate the previously recorded value or values, if no change has taken place in the unit.

Statistical Experiment

In a statistical experiment, the essential features of a random experiment and an engineering experiment are combined. In different statistical experiments, the importance of the random selection of a unit and its engineering measurement are separate features and must not be confused. The mere combination of these two types of experiments gives rise to such new terms as (1) trial of the statistical experiment, (2) random variable, chance variable, or stochastic variable, (3) sample space of values of a random variable for a particular statistical experiment, (4) population of values of a random variable, etc. Because of space limitations on this paper, we shall not write precise general definitions of all these terms. However, when these terms for a particular statistical experiment are understood and/or accurately defined, it is said to be a "well defined" statistical experiment. The reader can reflect upon these terms relative to the following experiments with the view to making them "well defined" statistical experiments in his own mind:

1. Toss a given thumbtack and observe whether the point comes to rest in the up or down (side) position.

2. Toss a given thumbtack one hundred times and count the number of times the point comes to rest in the up position.

3. Select randomly ten thumbtacks from a box, toss them all at one time, and count the number of thumbtacks coming to rest with the point in the up position.

4. Select randomly from inventory twenty-five missile components of a given type, operate them simultaneously and continuously to failure in a specified environment, observe the time-to-failure for each component, and compute the sample mean-time-to-failure for this sample.

5. Select randomly a missile system from the nation's supply of a particular type, operate this missile continuously in a specified environment by making immediate replacement of each failed component as failures occur, and at a given time count the number of failed components that have accumulated.

There are other new terms that arise from sampling distributions from "well defined" statistical experiments. However, we shall not give formal definition here but the chart of Fig. 1 explains some basic concepts.

BASIC PROBABILITY AND STATISTICAL CONCEPTS

The definition and meaning of the word probability has been the subject of much discussion for several centuries. In this paper, we shall attempt to make valid use of the von Mises [7] approach. Probability refers to the happening of an event which can occur as the result of a trial of a statistical experiment. The event in question can be a simple event, which corresponds to a single point in the sample space, or it can be a compound event, which corresponds to more than one point in the sample space of the statistical experiment. The terms mutually exclusive and exhaustive events are understood by most readers, and will not be explained in detail here.

There are two exceedingly important technical problems associated with probability: (1) the problem of determining a quantitative value for the probability of a particular event of interest, and (2) the problem of interpreting the practical meaning of this quantitative value, when it is given. How could one determine and interpret the probability value for the events of the following questions: (1) What is the probability that air lines pilot Frank House will have an accident on one of his regular flights?, and (2) what is the probability that fashion model Mary Jones will eventually get married? One can readily see that the probability in the first question can be estimated by statistical experiment but that the second question cannot. Let us consider event A which can result from a trial of a statistical experiment with a finite or an infinite number of outcomes or points in its sample space.

Success Ratio

Let n trials of a statistical experiment be made, during which the number n(A) of these trials have resulted in event A. Then, at any stage of experimentation, we can define the success ratio of event A to be

$$S_n(A) = \frac{n(A)}{n}, \qquad n = 1,2,3,\ldots .$$

Probability Limit Procedure

The probability limit procedure for event A is that of performing consecutive trials of the statistical experiment and observing the successive values of the success ratio of event A. If the sequence of observed values of the success ratio approaches a constant value π * as a limit as the number of trials

* This value π should not be confused with the irrational number 3.14159.... .

POPULATION	SAMPLE
1. The distribution of the N variable values x: $x_1, \ldots, x_N$ in the population is shown by the population distribution table, graph (dot diagram or histogram) or by a population distribution function fitted to these values. These population distribution characteristics are usually not completely known to the investigator.	1. The distribution of the values x: $x_1, \ldots, x_n$, $n < N$, in a random sample of sign^n is shown by the sample distribution table, graph (dot diagram or histogram) or by a distribution function fitted to these random sample values. These distribution characteristics are used to give insight to the the corresponding population distribution characteristics.
2. The cumulative distribution of the variable values in the population is exhibited by the population cumulative distribution table or graph, or by a cumulative distribution function, which are all usually unknown to the investigator.	2. The cumulative distribution of the values in a random sample is shown by the sample cumulative distribution table or graph or by a sample cumulative distribution function. These are used to estimate the corresponding population properties.
3. The distribution of values in the population is described by certain constants, called population parameters: a) The population mean is $\mu_x = \frac{1}{N} \sum_{i=1}^{N} x_i,$ b) The population root-mean-square or the population standard deviation is $\sigma_x = +\sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \mu_x)^2}$ These parameters are usually unknown to the investigator.	3. The distribution of the values in a random sample is described by the certain constants called sample statistics: a) The sample mean is $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i,$ b) The sample standard deviation is $s_x = +\sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2}$ These sample statistics vary from sample to sample. They are used to estimate the corresponding population parameter.

FIGURE 1. The Population and Sample Concepts in Statistical Experiments.

n increases without bound, then the probability of event A is defined to be the value π, and we write

$$P_r(A) = \operatorname{Plim}_{n\to\infty} S_n(A) = \operatorname{Plim}_{n\to\infty} \frac{n(A)}{n} = \pi.$$

Statistical Point Estimation

Let us consider again event A which can result from a trial of a statistical experiment with a finite or an infinite number of outcomes or points in its sample space. In statistical point estimation, it is implicit that the probability of event A is estimated from observed results of repeated trials of the basic statistical experiment. Thus, if n independent trials of the statistical experiment have resulted n(A) times in event A, then at this stage of experimentation, the success ratio is defined as the statistical point estimate of the value of the probability of event A and we write,

$$P_r(A) \simeq S_n(A) = \frac{n(A)}{n} .$$

A Priori Point Estimation

Let us consider again event A which can result from a trial of a statistical experiment with a finite or infinite number of outcomes or points in its sample space. In a priori point estimation, it is implicit that the probability of event A is estimated by a method not making use of the results of statistical experiments. Methods of a priori point estimation are often based upon enumeration of points in the sample space of the hypothetical statistical experiment or upon certain measures associated with subsets of values of the sample space.

In classical probability books, one will find a number of definitions of probability based upon what we are calling here a priori point estimation of the probability of event A. These definitions are usually attributed to Laplace. The reader might profit by reflecting upon the terms statistical experiment, random variable, trial, sample point, sample space, event A, probability of event A, statistical point estimate, and a priori point estimate of the probability of the events in the five statistical experiments mentioned above. The two questions about determining and interpreting the probabilities of an accident for Frank House and marriage for Mary Jones should be clearly understood. The reader may also find it useful to think of additional types of statistical experiments and the associated meaning of these various terms. A clear understanding of these statistical principles is essential for using statistical experiments to make decisions in the face of uncertainty.

STATISTICAL PROCEDURES FOR CONCLUSIONS WITH UNCERTAINTY

In the eight sources of uncertainty mentioned above, the questions involved the concepts of estimation or decisions in the face of uncertainty based upon the results of statistical experimentation. In the first seven of these questions, the various sources of uncertainty are difficult indeed to study. In coping with them, the engineer must resort to his best judgment and common sense in planning and executing an experimental program to get results which validly represent the desired future equipment. While much can be done to get helpful results from such experiments for future equipment, it is often impossible to place numerical values on the degree of confidence or on the risks of wrong decisions. However, in the eighth question about uncertainty due to randomness alone, the engineer can often do a great deal more, and in addition he can place numerical values on the degree of confidence in his estimates and risks of wrong decisions. There are two types of uncertainties contained in the factor of randomness: (1) uncertainty due to random selection of items for test from the population and (2) uncertainty due to not knowing which law of randomness or distribution function applies in a particular investigation [2].

Statistical Interval Estimation Procedure

This is a procedure for making an interval estimate of an unknown population parameter on the basis of a sample of fixed size, selected randomly from the total population under study and tested with proper instrumentation under the required conditions. The problem is to find limits to determine an acceptably short interval for which one can state with a fixed degree of confidence that the unknown population parameter is contained in the interval. These limits are called confidence limits and the interval, a confidence interval.

Statistical Decision Procedure

This is a procedure for making a decision about an unknown population parameter on the basis of a sample of fixed size, selected randomly from the total population under study and tested with proper instrumentation under the required conditions. The problem is to accept or reject a stated hypothesis about the unknown population parameter with a fixed probability of rejecting the hypothesis when it is true and an acceptably small probability of accepting

the hypothesis when it is false. This decision procedure was first formulated by Neyman and Pearson.

The following seven steps constitute a guide for carrying out the statistical estimation or decision procedures:

1. Define the basic statistical experiment, trial, unit, random variable x, population, parameters and other fundamental terms.

2. Assume distribution function or law relating to random values of x and the population parameters or other appropriate facts about the population of values resulting from the statistical experiment.

3. State statistical problem to be solved on the basis of a random sample of values x: $x_1, x_2, \ldots, x_n$ drawn from the population.

4. Fix appropriate values for the risks, confidence or sample size n.

5. Choose the statistic and state its distribution function under the above conditions.

6. Find critical values of the statistic for use with the statistical estimate or the decision.

7. Perform statistical experiment and make the actual statistical estimate or decision.

The statistical decision procedure for testing a hypothesis H_0 against an alternative hypothesis H_1 can be analyzed by an interesting chart showing the four state-decision combinations as shown in Fig. 2.

		DECISION MADE	
		H_0	H_1
STATE IN POPULATION	H_0	CORRECT	TYPE I ERROR
	H_1	TYPE II ERROR	CORRECT

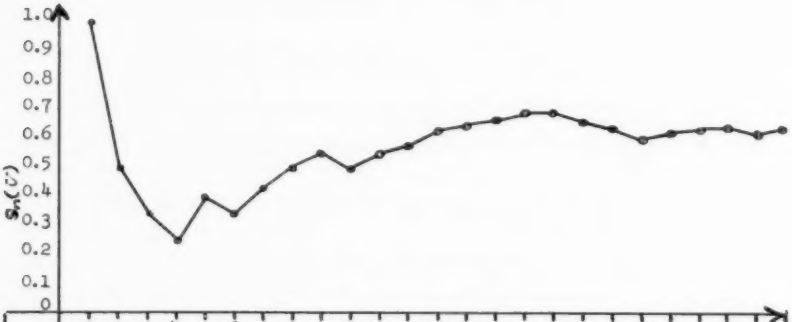
FIGURE 2. The State Decision Combinations Possible in a Statistical Decision Procedure.

STANDARD STATISTICAL ESTIMATION AND DECISION PROCEDURES AND THE EFFECT OF SAMPLE SIZES

The standard statistical procedures are those of point estimation, interval estimation and test of hypothesis about unknown population parameters. Let us study briefly the effect of the sample size upon these statistical procedures.

Effect of Sample Size Upon Point Estimation

The success ratio is a statistical point estimate of the probability of an event. Let us consider the thumbtack mentioned above and use a statistical experiment to make a point estimate of the probability of the thumbtack falling from a toss in the up position, U. A thumbtack was actually tossed twenty-five times from a leather container and at each trial the positions of up U or down D were observed as indicated in Fig. 3. The corresponding success ratios $S_n(U)$ can be



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	U	D	D	D	U	D	U	U	U	D	U	U	U	U	U	U	U	D	D	D	U	U	D	U
n(U)	1	1	1	1	2	2	3	4	5	5	6	7	8	9	10	11	12	12	12	12	13	14	14	15
$S_n(U)$	1	.50	.33	.25	.40	.33	.43	.50	.55	.50	.55	.58	.62	.64	.66	.69	.70	.66	.63	.60	.62	.64	.61	.62

FIGURE 3. The Effect of Sample Size Upon the Success Ratio of the Up Position U for the Thumbtack Experiment.

computed and plotted. This process is a probability limit procedure.

The success ratio $S_n(U)$ at any stage n is the statistical point estimate of the constant probability $P_c(U)$ that the thumbtack will fall in an up U position. This long run approach of the values of the success ratio to the value of the "so called" true probability is apparent. One might conclude that the sample size n at any stage is large enough by using several criteria such as (1) a measure of the rapidity of the approach of the graph of $S_n(U)$ to a horizontal straight line or (2) the maximum difference between the last several high and low points of the graph. The reader should think through this experiment and perhaps try the experiment with a thumbtack selected by himself. He might even be interested to compare this experiment with that of flipping a coin. Moreover, he might begin to reflect how one could study the effect of the length of the shaft or spike upon the probability of the thumbtack falling in the up position. It is likely also that the mathematically minded engineer would call the thumbtack a generalized coin.

Let us consider briefly the a priori estimate of the probability of the up position U in the statistical experiment on the thumbtack. This value is 0.50 since there are two points in the sample space of the experiment with one of

these points representing the up position U. Since there is no statistical experimentation involved in making the a priori estimate, the sample size can have no effect upon the value of the estimate.

Sample No.	Pop. Values x	Samples of Two		Samples of Three		Samples of Four		Samples of Five	
		Observations	Sample Means	Observations	Sample Means	Observations	Sample Means	Observations	Sample Means
1		2,4	3.0	2,4,7	4.33	2,4,7,8	5.25	2,4,7,8,14	7
2	2	2,7	4.5	2,4,8	4.67	2,4,7,14	6.75		
3		2,8	5.0	2,4,14	6.67	2,4,8,14	7.00		
4	4	2,14	8.0	2,7,8	5.67	2,7,8,14	7.75		
5		4,7	5.5	2,7,14	7.67	4,7,8,14	8.25		
6	7	4,8	6.0	2,8,14	8.00				
7		4,14	9.0	4,7,8	6.33				
8	8	7,8	7.5	4,7,14	8.33				
9		7,14	10.5	4,8,14	8.67				
10	14	8,14	11.0	7,8,14	9.67				
	35		70		70		35		7
	$\mu_x=\frac{35}{5}=7$		$\mu_{\bar{x}}=\frac{70}{10}=7$		$\mu_{\bar{x}}=\frac{70}{10}=7$		$\mu_{\bar{x}}=\frac{35}{5}=7$		$\mu_{\bar{x}}=\frac{7}{1}=7$
	$\sigma_x=4.1$		$\sigma_{\bar{x}}=2.51$		$\sigma_{\bar{x}}=1.68$		$\sigma_{\bar{x}}=1.03$		$\sigma_{\bar{x}}=0$
	$R_x=12$		$R_{\bar{x}}=8$		$R_{\bar{x}}=6$		$R_{\bar{x}}=3$		$R_{\bar{x}}=0$

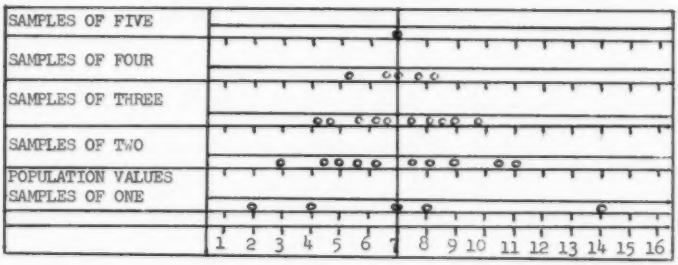

SAMPLES OF FIVE
SAMPLES OF FOUR
SAMPLES OF THREE
SAMPLES OF TWO
POPULATION VALUES
SAMPLES OF ONE
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

FIGURE 4. The Effect of Sample Size Upon the Sample Mean as a Statistical Point Estimate of the Population Mean.

The sample mean $\bar{x}$ for any sample size is often used as a statistical point estimate of the population mean μ_x. In Fig. 4 a population of five values is shown together with all possible samples of two, three, four and five values each. The numerical values of the various sample means are also shown in the several columns and are plotted in the second part of the figure. The population standard deviation $\sigma_{\bar{x}}$ and the population range $R_{\bar{x}}$ for each distribution of $\bar{x}$'s are shown in the last two rows of the chart.

For each sample size, we observe that the population mean $\mu_{\bar{x}}$ of all possible sample means $\bar{x}$ is equal to 7. This highly desirable property expresses that the sample mean $\bar{x}$, for any sample size, is an unbiased estimate of the population mean μ_x. We observe also from the graph that as the sample size n is increased the individual sample mean values cluster closer to the population mean. This property can also be seen by looking at the values of the population standard deviation or range in the last two rows of the chart.

CONCLUSION

The properties of unbiasedness and cluster of the sample mean $\bar{x}$ as an estimate of the population mean μ_x lead in statistical theory to many criteria for answering the question "When is a small sample large enough?" The criteria are often based (1) upon consideration of length of the confidence interval in estimation or (2) upon certain specified requirements placed on the operating characteristic curve of a decision procedure in the face of uncertainty. The reader can find additional interesting discussions of these principles in [4] and [6].

REFERENCES

[1] Bowker, Albert H. and Gerald J. Lieberman, Engineering Statistics, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1959.

[2] Chernoff, Herman and Lincoln E. Moses, Elementary Decision Theory, John Wiley and Sons, Inc., New York, 1959.

[3] Feller, William, An Introduction to Probability Theory and Its Applications, Second Edition, John Wiley & Sons, Inc., New York, 1957.

[4] Fraser, D.A.S., Statistics: An Introduction, John Wiley & Sons, Inc., New York, 1958.

[5] Mood, Alexander M., Introduction to the Theory of Statistics, McGraw-Hill Book Company, Inc., New York, 1950.

[6] Thrall, R.M., C.H. Coombs, and R.L. Davis, Decision Processes, John Wiley and Sons, Inc., New York, 1954.

[7] von Mises, Richard, Probability, Statistics and Truth, Second revised English edition prepared by Hilda Geiringer, The MacMillan Company, New York, 1957.

HOW TO PLAN AND IMPLEMENT AN INCOMING MATERIAL CONTROL PROGRAM

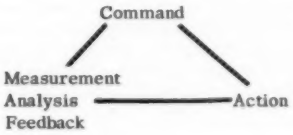
C. Gadzinski
Raytheon Company

Introduction

In this paper we shall describe the work, organization and measures of work for Incoming Material Control. We shall develop an Incoming Material Plan through the first five years of a hypothetical new company. One reason for this procedure is to permit those who have varying degrees of personal maturity as well as organizational sophistication to recognize their problems and thereby gain some insight toward their solution. Another reason is to illustrate that organizations are not built overnight nor are results produced effortlessly.

Principles of Control

Consider the concepts of command, action, measurement analysis and feedback. These may be described as follows:



and can be referred to as components of a closed loop system. This is another way of saying direction, execution and evaluation of results. These as you will recognize, are the principle tasks of management and incidentally, the basic elements for the control of any operation. It is upon these principles that we will develop our plan for Incoming Material Control. Before we do this, however, we must have a clear conception of the work that needs to be done and some idea as to what measures should be established to assure that the work is being successfully accomplished.

The Work of Incoming Material Control

We can say that the task of Incoming Material Control is to assure that what is purchased meets the purchase order requirements. This, however, has as an implicit assumption that the Purchase Order is complete; that the drawings are complete and all referenced data pertinent to the purchase order were sent to the vendor and that the vendor understood the requirements plus had the ability to perform them. All of these are rather rash assumptions.

If, on the other hand, we say that the task of Incoming Material Control is to assure that product purchased meets the end use requirements of the equipment for which it is purchased, then Incoming Material Control must concern itself with the composition of the purchase order, the completeness of the drawings, the referenced data supplied to the vendor and the vendor's understanding and ability to comply with the requirements set forth in the purchase agreement.

Organization for Incoming Material Control

We will organize the Incoming Material Control function into a Planning Group and a Doing Group. The Planning Group will do the Quality Control Engineering work and

will provide the technical direction for the Doing Group which will perform the Inspection and Test. The reason for doing this is to separate the thinkers from the doers. Experience indicates trying to combine the two, results in neither good thinking nor effective doing.

The Work of Quality Control Engineering - Incoming Material Control

1. Establish a system of communication including feedback system for data.
2. Review specifications to assure conformance to end use requirements and establish visual standards.
3. Classify characteristics.
4. Establish Quality Levels.
5. Plan the work.
6. Establish an analysis plan for the data obtained.
7. Diagnose problems and secure corrective action.

Let us consider each of these elements in detail:

1. Establish a System of Communication Including Feedback System for Data
   Unless you know your inputs (that which describes work to be performed) and outputs (measurable results of work accomplished), you don't know your job. To explain this further in the area of Incoming Material Control, you must have some notice that material is being ordered so that you can prepare your tests and inspection plans in order to be ready for the material when it is received. If we consider that the control of vendor quality is directly related to the quality of information sent to the vendor, it becomes apparent that Quality Control must exercise some control over this factor. See Figure 1.

2. Review Specifications to Assure Conformance to End Use Requirements and for Such Things as:
   (a) paint finishes.
   (b) nicks, scrapes, dents and surface finish.
   (c) wire dress, layout of components and cables.
   (d) burrs, rough edges and pin holes.
   (e) roundness, flatness, concentricity, parallelism etc.

   This work is not to be confused with the work of design engineering or industrial engineering, i.e. Quality Control Engineering is not to specify the paint or the method of application, unless this is the only way to assure conformance to requirements, but it is to establish a standard. This standard will be used to compare that which has been produced against that which was required. Quality Control must make sure that the people who are trying to work to the standard understand it. Similarly, pictures of approved soldering, wiring or cabling processes should be produced so that everyone knows what is acceptable wiring workmanship quality. Roundness, flatness, concentricity, parallelism should also be specified in such a fashion as to avoid misunderstanding.

3. Classification of Characteristics
   Identify what is important based upon an understanding of where each component fits into the next higher assembly.
   A simple example: -

FIGURE 2.

EXAMPLE OF FLOW PLAN FOR SIMPLE ORDERING PROCESS

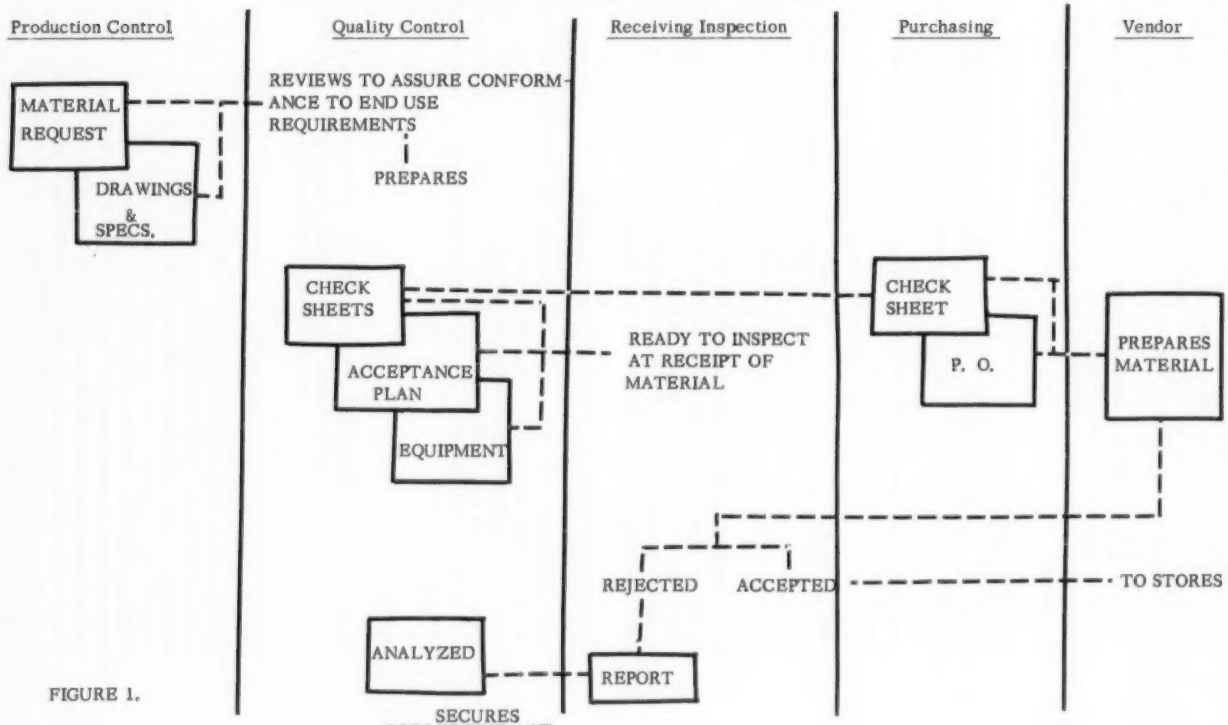


FIGURE 1.

The functions of Parts A and B are to simply go together as shown. Parts A and B must have several other dimensions in order to describe their general configuration; however, the important function is that the two parts go together; hence, dimensions C and D are critical.

Classification of characteristics is a method of describing the relative importance of the various characteristics of a part, component or assembly based upon an understanding of the performance required of that piece of equipment.

4. Establish Quality Levels
Quality Levels are established on each part and are based upon the performance required and upon the economy of 100% inspection and test versus sampling inspection and risk taking. Since a chain is only as strong as its weakest link, so too, any assembly is only as strong as the members which make it, hence, the evaluation of the performance requirement of each part is a vital task for Quality Control.

5. Plan the Work
Describe who is to inspect what, where, when, at vendor's plant or at Receiving Inspection, and how. Describe sample (amount), frequency, type and disposition. Provide equipment for performing the Inspection and Test.

6. Establish an Analysis Plan for Data Obtained
An analysis plan for the data should exist prior to the request for the data. This analysis may be in the form of frequency distribution, correlation, rating, or simply review; however, it should be purposeful and the purpose should be clearly stated and the consequences of it understood.

7. Diagnosing Problems and Securing Corrective Action
The Quality Control Engineer must diagnose the quality problems and identify their originators, i.e. Purchasing, Engineering, Industrial Engineering etc., and present the problems and their consequences in such a fashion that a reasonable individual will accept them and commit himself to their solution by a given date. All men want to do a good job. They just want to know where they are not doing too well and to be given an opportunity to correct their own shortcomings.

Example: -

Problem: - Drawing not calling for proper hardware on large sheet metal cabinet. Result: - Each cabinet as received must be torn down and all hardware replaced.

Originator of Problem: - Engineering

Quality Control Engineer's Presentation: - There is a minor error in the drawing whereby the hardware specified is of a different material from the base metal. This is contrary to the requirement of MIL E 5400, hence all material being received is being rejected and because of production necessity, is being reworked. As I have mentioned, we need these parts. Each cabinet requires eight hours of rework. You can imagine the frustration of Purchasing and Production Control in having expedited these units only to find them rejected. I've already spoken to Purchasing and they have stopped delivery from the vendors. It's up to you. When can I expect the drawing change?

Engineer: - I'll get on it right away.

In summary, the work of Quality Control Engineering - Incoming Material Control is as follows:

1. To establish a system of communication with affected organizations of manufacturing and provide a system of feedback and analysis of inspection and test data in order to diagnose problems and secure corrective action.

2. Review specifications to assure conformance to end use requirements and establish workmanship standards.

3. Classify important characteristics based upon performance requirements.

4. Establish Quality Levels both with respect to performance required and economy of 100% inspection and test versus sampling inspection and risk taking.

5. Plan the work and provide the equipment required for inspection and test.

6. Establish an analysis plan for the data obtained.

7. Diagnose problems and secure corrective action.

Measures for the Control of Incoming Material Control

What is the cost of Incoming Material Control versus the cost of Purchased Material? This is the usual benchmark from which all planning for Incoming Material Control begins. Unfortunately, no standard exists for this ratio. A survey of business in a given geographic area showed this to be anywhere from 3 to 30%. In a business which is highly complex and whose purchasing volume is 60% of the sales billed, the goal of 3% as the cost of labor has been established. This includes all labor for Incoming Material Control.

Test and Inspection equipment ratios are meaningless since the Test and Inspection must be planned for the job.

Space for a facility which would require various types of environmental equipment would be approximately 200 sq. ft. per man excluding the space of aisles. This figure for space required varies between 130 sq. ft. per man to 250 sq. ft. per man. It considers Receiving Inspection only. It does not consider the combined space of Receiving and Receiving Inspection.

Other measures to be considered are as follows:

(a) a measure of the rate of rejection at Receiving Inspection.
(b) a measure as to the disposition of rejected material in terms of:
    1. Return to vendor.
    2. Rework.
    3. Use as is.
    4. Scrap.
(c) a measure to show the concentrations of rejects by vendor.
(d) a measure of the time in processing material through Receiving Inspection.
(e) a measure of the productivity of Receiving Inspection in terms of lots of material handled by each inspector.
(f) a measure of the distribution of time required in processing material through the disposition routine in order to determine whether material should be returned to a vendor, reworked, scrapped or used as is.
(g) a measure of the backlog and distribution of time of the material in the hold area of Receiving Inspection.
(h) a measure of the backlog of material to be inspected at Receiving Inspection.
(i) a measure of the productivity of the quality planning function.

Five Year Plan for a Hypothetical New Company

Now that we have developed the work that needs to be performed in Incoming Material Control and some initial measures that would be useful to control this effort, let us consider a five year plan for the development of an Incoming Material Control Program. This program will be developed assuming that we are beginning a new company. An organization plan for a business just starting is described in Figure 3.

As targets for the plan, we will consider cost of operation, as a percent of value of material purchased, reject rate percent, time for processing material through Receiving Inspection, and productivity measured in lots of material handled per man.

Our goals may be described as follows:

	YEARS				
	1	2	3	4	5
Cost of Operation	7%	5	3	3	3
Reject Rate %	25%	15%	5	5	3
Time	10 days	5	3	2	1
Productivity	15/lots per man	20	25	30	35

In the first year, our primary effort will be to establish a basic capability in people and equipment. We shall not consider much else. The problem of securing qualified people and getting them to work together will be so great as to preclude any excursions into other considerations. For failure to observe this simple rule many of our Quality Control programs get into trouble. There's not much sense in attempting to tell Purchasing how to select and control their vendors if you do not have an experienced hard hitting organization to follow through. To attempt to tell other people how to run their business without having your house in order is courting disaster.

If we have an organization plan comprised of a planning function and doing function, the problem of roganization will be those of getting the Planning Group to do the Planning and the Doing Group to do the doing. See Figure 3. This I assure you is no mean task. The doers will attempt to do the planning due to the inexperience of the planners and the planners will attempt to do the doing because of the inexperience of the doers.

To train the planner to do an adequate planning job, and the doers to follow the plan will be a job that will call for the utmost of tact and persuasion.

Some policy decisions will have to be made with respect to labor grades. This will have to be established in order to control the direct labor mix and provide a guide to the planning function as to the degree of detail they are to pursue in their planning. It's obvious that the more detailed or pre-planned the plan, the less skill required to execute it. Also significant at this time will be the training value of the plan. It must be considered that although the plan may begin to be detailed for purposes of training at some time the need for this training will cease or at least be diminished.

In a starting operation with a great many new people, their selection and performance appraisal are of primary importance. The problems of selection can fill volumes. However a suggestion to attempt to balance experience with potential if successfully accomplished will provide many returns. In the area of performance appraisal, the

MANAGER
QUALITY CONTROL
ADMINISTRATION AND PROCEDURES
INCOMING MATERIAL CONTROL
IN PROCESS CONTROL
FINAL TEST
EQUIPMENT DESIGN
Planning
Doing
Planning
Doing
Planning
Doing
Planning
Doing
Planning
Doing

FIGURE 3.

importance of making sure that the appraisal is correct and just and that the same standards of performance are used as will be used a year or two years hence cannot be over-stressed. It is most difficult to discipline someone who has had a series of satisfactory ratings when the fact of the matter is that the appraisor lacked the courage to present the facts.

A report in Dun's Review mentioned that the top executives of the country's leading industries could make the proper selection of the right man for the right job in only four out of five cases. This would indicate that for an organization of 100 men, 20 would be misfits or marginal performers. It is the responsibility of the manager to identify these people and to see to it that they are either removed from the payroll or placed where they can become more effective.

In addition to the problems of staffing, the manager will face the problem of equipping his operation and performing the work required. In this respect, perhaps a few guide rules will help, namely:

1. Identify major equipment problems and establish control. This would be on those items where because of size, value, cost or safety considerations, special effort for control will have to be made.
2. Establish operational capability before attempting to help somebody else.
3. Anticipate that not all of the major equipment problems will be identified at the first try; some others will come up that must be dealt with.
4. Wait for results of performance on a given plan before attempting to apply effort at improving or changing it.
5. Provide a plan to identify the chronic few problems that will occur, to prevent going after the sporadic many.
6. Evaluate each problem as to whether it occurred for lack of a procedure to anticipate it or because of a lack of competence of the people employed to try to solve it.

The tasks of getting operational and controlling, motivating and integrating his people are the principal tasks of a manager in his first year of operation. As mentioned previously the anticipated performance for the first year was as follows:

Cost of Operation	7% of value of purchased material
Reject Rate %	25%
Time	10 days
Productivity	15 lots/man

In preparing for the second year we must measure the first year's performance against what was anticipated. The degree to which the plan was fulfilled will determine the amount of effort required to get to the second year's objectives. Let's suppose that we're right on schedule. Hence, our second year's objective will become as follows:

Cost of Operation	5% of value of purchased material
Reject Rate %	15%
Time	8 days
Productivity	20 lots/man.

These objectives can be accomplished through studying vendor performance, paperwork simplification, area utilization, equipment utilization and manpower productivity. Manpower productivity includes the doing manpower as well as the planning manpower. In looking at the planning manpower, what time is spent putting out fires and what time is spent trying to prevent them. In looking at doing manpower, what is productive inspection and test and how much of the inspection and test time is time wasted in just getting set up to test or inspect, i.e. what time is spent chasing prints or tools and

getting instructions, or wrapping and unwrapping packages? A work sampling study will provide this information quite readily.

Once this information is accumulated, an attack on the problem of increasing productivity can be started. The methods of industrial engineering can be employed to bring about certain improvements. However, they will only produce limited results when considered within the basic conception of Incoming Material Control, i.e. as long as we begin from the premise that Incoming Material Control is Receiving Inspection, then we can only go so far in improving Receiving Inspection before we face the fact that the cost of further improvement would outweigh the savings. Hence, a new approach must be sought. Vendor Certification wherein we utilize vendor inspection and test data to reduce the inspection and test required at Receiving Inspection is one of these approaches.

Vendor Certification does not mean sending uninstructed inspectors out into the vendors' plants to attempt to perform an inspection. This is clearly the blind leading the blind. Vendor Certification is an effort which has engineering direction. The substance of this program would be to influence vendor selection, to assure the control of information going to the vendor, the assuring that the vendor received and fully understood the information and the monitoring to assure the vendors' conformance to the requirements. This work would be accomplished through planned visitations to the vendor.

It must be kept in mind that the only legal contact with the vendor is through Purchasing. Special care must be exercised to assure that all contacts with the vendor are made in conjunction with or through Purchasing and that all agreements are confirmed by Purchasing through an amendment to the Purchase Contract.

In planning the vendor certification effort it is helpful to divide your vendors in the following categories:

1. Subcontract Vendors - Vendors who have been given a work statement and who are to design, develop and fabricate a given piece of hardware. These are usually the high dollar volume items.

2. Contract Manufacturers - Sheet metal and machined parts fabricators who perform work to drawings and specifications supplied by the customer.
3. Casting Manufacturers.
4. Suppliers of AN & M.S. Parts.
5. Suppliers of Basic Electrical Components.
6. Suppliers of Raw Material, Paints, Chemicals and Lacquers.

Experience indicates that 50% of the effort of Receiving Inspection is on the product of categories 1 and 2. Hence it is in these product areas that the greater effort should be made.

The planning for vendor certification will include the selection of vendors, the establishing of procedures and relations with Purchasing and Production Control, the preparation of specifications and instruction books and perhaps, the planning of a seminar to acquaint the people in the plant as well as the vendors with the program. One vital consideration will have to be the amount and rate of applied effort and the consequent rate of results. Obviously, the greater the applied effort the quicker the result. This would, however, have to be balanced with respect to people available, time and money.

A reasonable program is suggested in Figure 4. (See following page).

VENDOR CERTIFICATION ACCOMPLISHMENTS EXPECTED

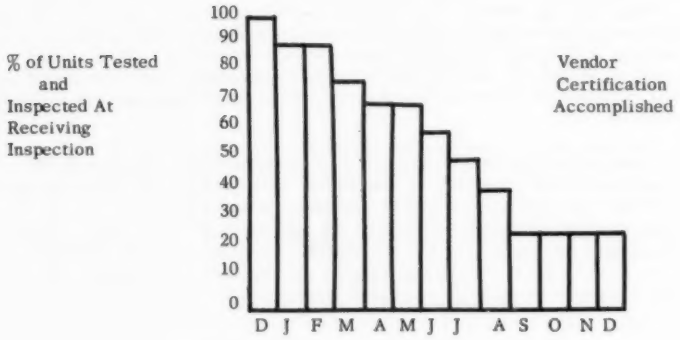


FIGURE 4.

At the end of the second year, we would once again review our performance against the plan and thereby determine what would be appropriate future conduct. Once again, if we assume that our objectives have been fulfilled, we can proceeed with our plan in the third year. Our objectives for this period are:

Cost of Operation	3% of value of purchased material.
Reject Rate %	5%
Time	4 days
Productivity	25 lots/man

The third year will be a year devoted to the implementation of the Vendor Certification Program.

It would be primarily through the efforts of vendor certification that the third year objectives would be achieved.

Cost of Operation	3% of value of purchased material.
Reject Rate %	5%
Time	2 days
Productivity	30 lots/man

The plan for accomplishing these objectives would be through the refinement of the Vendor Certification Program, the introduction of semi-automated equipment and through methods simplification.

The fifth year's objectives of:

Cost of Operation	3%
Reject Rate %	3%
Time	1 day
Productivity	35 lots/man

would come about by the force of the accumulated experience.

Conclusion

We have defined the work of Incoming Material Control and some measures for assuring the work has been accomplished. We have developed a dynamic organization plan over a period of time to illustrate that organizations are not built over night nor are the results produced effortlessly.

In conclusion, we might observe that the mission of the manager is to make things happen. The things that happen must occur, however, by plan. To make the plan work, the manager needs people. To the extent that he can select, instruct, motivate and control his people to produce the desired results therein will lie his genius.

STATISTICAL QUALITY CONTROL APPLICATIONS IN THE FOOD INDUSTRY

C. B. Way
Green Giant Company

In recent years, there has been a tremendous advance in the use of quality control techniques based on mathematical applications of the laws of probability. It has become quite apparent that 100% inspection and testing of a production line is seldom feasible, particularly when the test is destructive; i.e., the testing of ballistic shells. Systems for estimating and controlling the quality level of production from measurements of small samples have been developed and it is possible to estimate the degree of precision of these estimates from the data. Similar sampling techniques for estimating the quality of given lots of manufactured items or raw materials with known degrees of risk have also been developed. The use of these various techniques is what we know today as "statistical quality control." The interest in their applications is evidenced by the ever increasing membership in the American Society for Quality Control, now numbering in the neighborhood of 12,000.

"Heralded as one of the great new developments of industrial technique," says Dr. Kramer, (12) "Statistical quality control has been thought by many to be primarily related to such uniform product endeavors as making tools, machines, and the like.

"Thus, application of SQC in biological fields, such as food production, has been presumed less feasible, particularly because here difficulties abound due to inherently greater variability of the raw materials.

"But this is a restricted viewpoint. Actually, much of the early work in SQC was done in the food field--precisely because natural variability of the materials was so great, and factors effecting quality of product so many, that better means were required to distinguish cause and effect and thus speed the way to remedial measures through changes in material, equipment, or production procedures."

Kornetsky (15) suggests, "Variability that exists in food materials has been a deterrent in the application of statistical quality control charts to food manufacturing processes. While it is recognized that food products are more variable than the manufactured articles of other industries, this in itself does not exclude the use of control charts; it merely indicates that control (limits) must be wide and must be adjusted to the situation."

In 1955, Dr. Hosking (8) reported the results of a survey of the use of SQC in the food industries. That report indicated that 64% of all food industries were using SQC. Dr. Stier (27) makes the statement, "As food processing has become more and more a continuous process and less of a batch process, the use of statistical control techniques has increased."

The purpose of this paper is to describe very briefly some of these applications. No attempt has been made to detail the methods used nor is the list complete. Many of the references do describe individual techniques in detail and can be used by the processor as applicable.

That the food industry recognizes the importance of statistical procedures is evidenced by the fact that the National Canners Association Laboratory Manual (34) has a very complete chapter on statistical procedures. This book describes variability, probability, control chart methods, sampling, correlation, some basic taste panel techniques and other procedures commonly used.

There are other training manuals such as those published by the University of Iowa (35) and the St. Louis (36) sections of the ASQC, as well as a number of good books on statistical methods and their application to quality control. One of the important values of the afore mentioned manuals lies in the fact that they are easily understood by those with little or no basic statistical background. One need not understand the complicated mathematical derivations to use and understand the methods themselves. The lack of realization of this fact has been a deterrent to the applications of SQC in many cases. It cannot be overstressed.

As these SQC procedures have come to the attention of food processors, they have looked about for possible points of application. One of the most prominent is that of fill of container which seems to be a straightforward and rewarding application. Much has been done and written on SQC applied to fill of container in the food industry. In 1954, the National Canners Association Convention included a symposium on fill(22). Some basic methods of fill control using $\bar{X}$ and R charts were explained and the advantages and improvements were demonstrated. The advantage of reducing the variation and thereby reducing the average fill without increasing the number of under-fills was demonstrated.

Twigg (30) gives a method for checking on the uniformity of individual filler pockets. He demonstrates by example that malfunctioning of a filler can cause non-uniform filling among the pockets which, if corrected, would save an estimated 2½% of the fill-in weight, this by reducing the average without increasing under-fills. He describes a method of constructing $\bar{X}$ and R charts and mentions some advantages of this type of control system, such as: (1) - Catching a gradual change in fill level which might otherwise go unnoticed. (2) - Preventing unnecessary and often harmful adjustments and changes. (3) - Locating assignable causes of variation.

Bremer (2) emphasizes the practical aspects of control chart work on a packaging line. He suggests that, for some operations, control limits based on two standard deviations rather than the conventional three standard deviations may be beneficial. Limits set up for individual packages rather than averages are often useful because the results more nearly approximate field practices of regulatory agencies. The use of a second set of control limits twice as wide as the regular control limits was attempted, but it was found that the operators were biased and tended to try to stay within the first and second lines on the high side, which nullifies much of the improvement of the control method. The performing of the weight checks in consistent communication with the operators made them realize the importance of the control of net weights. It is also pointed out that, in some cases, fillers or filling units produce such good weight patterns that the fill-check operation became monotonous because the weights were all either "0" or plus or minus 1. This was corrected by changing to a scale of increased sensitivity which produced less monotony and improved the operator's attention. It is emphasized that the function of Quality Control is a service and definitely not a police or punitive function.

Reynolds (23), "offers a simple yet effective method of establishing a practical, workable weight control system over most types of filling operations." He makes the very practical suggestion that whole numbers should be used rather than fractions. For instance, 2/32 over the scale zero should be recorded as +2. He suggests that increments of weighing should be small enough to show changes, but no smaller than necessary. For instance, if the normal expected range of weights varies over $\pm$ 1/4 lb., probably increments of 1/8 or 1/16 oz. are acceptable. If, on the other hand, the total spread of weights amounts to $\pm$ 1/16 oz., probably increments of 1/64 oz. would be needed. Both group and random distribution methods of capability testing of the machines are given. The group plan consists of drawing samples of five consecutive packages from the line and weighing as is done in normal $\bar{X}$ and R chart construction. The

random distribution plan consists of drawing one sample at a time and obtaining a single sample distribution. It is pointed out that the group plan should normally be used because it is not affected by changes in the over-all fill level. A method for establishing the average fill desired, based on a designed percentage of underweights, is given. This is simply done by setting the average fill as far above the minimum fill as necessary to yield the desired per cent of underweights as determined from the table of cumulative probabilities of the normal distribution curve. It is cautioned that in actual practice, the per cent underweights will apparently be somewhat less than the desired because of a peculiarity of recording, wherein, if the scale is being read to 1/8 oz., any weights which fall between 1/16 above and 1/16 below the minimum or "0" are counted as exact weights. Actually, a part of them would be shown to be below if the units of measurement were smaller.

Another source (31) presents a table of standard deviations of the fill of different food products. This demonstrates the direct relation between the ease of filling and the standard deviations obtained. In most cases, heterogeneous products have higher standard deviations of fill than the more or less homogeneous products; i.e., whole beets versus juices.

Harrison (7) explains the relation between individual package weight and case weight. It is shown that there will be very few (or no) cases underweight even though a small percentage of the individual packages are underweight, if the underweight packages are randomly distributed. On the other hand, if the underweight packages are clumped, as is normally the case, the percentage of cases which are underweight will approach that of the individual packages. It is suggested that some method of randomizing packages from a filling line be investigated so as to minimize the percentage of cases which are underweight.

Donohue (3) describes some of the problems involved in controlling a filling line which has several spouts or several machines and gives a solution involving a "variance analysis technique." This involves checking each individual machine and/or each spout for abnormal adjustment or variation. Analysis of the various interactions indicated, among other things, inherent variation in the stock fed to the machines. This work resulted in readjustment of some machines, redesign of the packaging line, the setting up of operating control limits and a system of individual machine checks and "realized a state of control that approaches the inherent variability of the machines."

Gieseker (6) uses a lot plot sampling procedure adapted to line control (fill, in this case). This is not as tight a control as $\bar{X}$ and R chart control. It is "practical when many of the assignable causes for variation - - are known from past experience."

Rowe (25) and McBride (21) have both shown that normal variation is recognized by the regulatory agencies. Regulations generally state in effect that the fill of a product must average a specified weight but with "reasonable variations consistent with good filling practices."

Weiss (33) describes how the Food and Drug Administration sets up standards based on what it believes to be good manufacturing practices. He also points out that "the use of a lower limit of acceptability constitutes a tolerance." "Assuming that the variability in the product is not controllable, if production were changed so that the average weight were to be shifted from the recommended fill to that just above the lower limit of acceptability, the end result would be that a substantial proportion of the product would be substandard. The limits are set so that by conforming to the recommended fill, the ensuing product will be considered acceptable. This is not a tolerance." He also describes a multiple sampling plan or modified sequential plan for checking on the conformance to standards. Thus, products of good character or very inferior character are acted upon quickly while those of border line character re-

quire more samples for final action.

The American Marketing Service (USDA) has developed multiple sampling plans for inspecting fruits and vegetables (1). These tend to reduce sample sizes (from the former sampling plan) and, thus, the cost of inspection.

Kornetsky (16) suggests the use of control charts for determining the effects of the sorting and trimming operations in sweet corn. Samples of the raw product are drawn and checked for defects and plotted on a control chart. Presumable, if they are out of control on the high side, the number of inspectors would have to be increased or the rate of processing slowed down in order to effectively inspect and trim the corn. Similarly, samples are taken after trimming and inspection and the defect infestation measured. If out of control, it would show a need for improving the inspection and trimming section. The use of a control chart for controlling the consistency of cream style corn is also demonstrated.

Kramer (14) demonstrates the relation of sample size to error in the sampling of raw green beans as well as suggesting a sampling schedule based on the Mil Std 105A. He shows the effect of sampling error on the per cent seed, shear press and per cent fiber determinations on the sample. In-plant control of the snipping and filling operations are briefly described and a control chart demonstrating finished product control is suggested. In another article (11) he points out several types of sampling procedures for food products.

Kramer and Steinmetz(13) obtained a correlation coefficient of .86 between the shear-press values of raw beans and the per cent fiber of canned beans. A table of shear-press versus per cent fiber is given and, from this, a system for raw produce grading is proposed, whereby beans below a given shear-press reading are generally acceptable and those above a given shear-press reading would probably produce a sub-standard canned product, the area between the two levels being questionable. Since green beans increase in size with maturity, this method lends itself to in-plant control of green bean processing based on the raw product shear-press values. For instance, with low shear-press values, one might can the 1-2-3-4 sieves separately or in various combinations as whole beans, and cut the large sieves. But, if the raw shear-press were high, one might discard the larger sieves without canning because there would be very little chance of their being better than sub-standard. (The relation of shear-press to fiber should be determined separately for each variety of beans.)

SHEAR-PRESS	PER CENT FIBER	
1200	0.01	
1250	0.03	
1300	0.05	
1350	0.07	
1400	0.09	
1450	0.11	Questionable
1500	0.13	Area
1550	0.15	
1600	0.18	
1650	0.22	
1700	0.27	
1750	0.35	
1800	0.46	
1850	0.65	
1900	1.00	

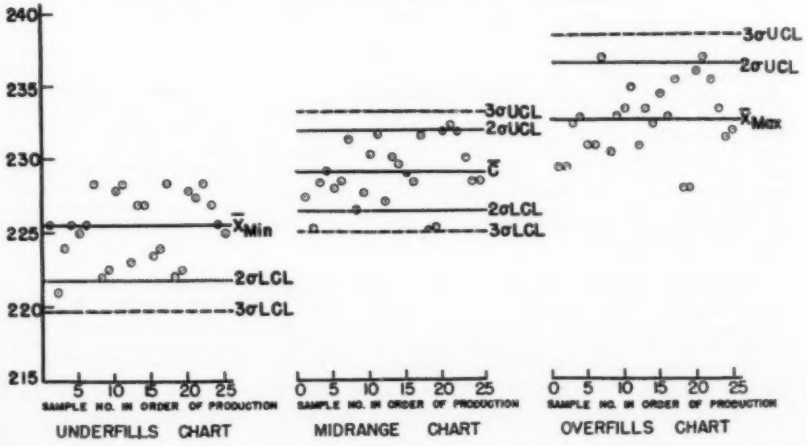
Statistical techniques have been employed to experiments involving the testing of cans which have considerably reduced the cost of these experiments. Normally, cans of varying structural designs are filled, processed and cooled in water highly contaminated with gas-forming bacteria

so that the cans which leak, will almost certainly swell in a short time. In at least one test of this nature, the difference in number of swells among the various types of can structures was insignificant because the number of cans tested was too small. Statistical calculations have greatly reduced the risk of this type of needless expense (since the test had to be re-run) by showing the number of cans which needed to be tested in order to determine any significant difference in the number of spoiled cans among the variables. The following table is the result of one such calculation:

Assume Control cans to have spoilage in cans per thousand	If it is necessary to determine differences between groups of ( ) cans per thousand	This number of cans should be tested
5.0	1	40,000
4.0	1	33,000
3.0	1	25,000
2.0	1	17,000
1.0	1	10,000
5.0	2	12,000
4.0	2	10,000
3.0	2	7,000
2.0	2	4,500

Tarver, et al, (28) describes a statistical method for selecting flavor panel members to be used for detecting small flavor differences in paired samples. They have very carefully pointed out that this type of panel is a tool for the research and development laboratory and cannot be considered as a consumer acceptance panel.

Tarver and Schenck (29) describe a method of using extreme value control charts to control the actual maximums and minimums rather than averages or ranges. The following charts show some sample data plotted as extreme value control charts. The chart on the left is obtained by plotting the lowest sample value from each subgroup. Likewise, that on the right is obtained by plotting the highest value from each subgroup. The average of the maximum and minimum values are plotted on the midrange chart. The major objective of these extreme value control charts is the specification of tolerances for over-fills and under-fills under normal operating conditions. Thus, in this example, any fill below 220 grams is an under-fill and any fill over 238 grams is an over-fill.



Weimerskirch (32) reports a method of top seam control of the can closing operation and an acceptance sampling plan for incoming cans. One of the results of this work was an increase in the control limits beyond the machine capabilities which had the advantage of reducing the amount of actual controlling. Such action, however, is not always feasible and should be taken only after careful consideration. The acceptance sampling plan was sequential in nature but truncated after a certain number of samples had been examined, if no previous decision had been reached. Different operating characteristic curves are used for each of two classes of defects. Considerably larger AQL and LTPD values are used for the less serious defects as is to be expected.

Lang (17) reports on the use of an acceptance sampling plan based on Mil Std 105A for glass jars. In this case, three classes of defects with appropriate AQL values are used.

Filice (5) describes some control chart applications to the fill of fruit, pH control, total solids of tomato products, and labeling lines. He demonstrates a very rewarding result of raw product research on pears. A system for determining maturity of a lot of freshly harvested pears was developed and from this, by continual sampling of these lots, a system for predicting the number of storage days necessary to ripen to optimum maturity for canning. In addition, lots of excess variation can be segregated and handled differently in the canning process.

McPhail (20) describes the use of quality control charts in controlling the acidity of vinegar with a resultant saving.

Ritchell (24) describes the problems involved in a packaging line such as torn labels, dog eared labels, dented cans, dirty cans, unsealed or poorly sealed cases, poor stenciling, etc., and the control system used to reduce the number of packaging defects. In this case, a "number defective" chart is used. A sample is drawn periodically after labeling, casing, and sealing and the defectives noted and plotted. Control limits were calculated from the first data taken in this manner and have dropped considerably since the inception of this packaging control program. It was found necessary to set up different limits for different types of labels, since foil labels are more difficult to handle than paper labels.

A number of cases describing problems and their solution through statistical quality control techniques were presented at a National Canners Association problem session in April 1956. Many of these have previously been mentioned at least by type; however, there are one or two which show a different approach to these problems and warrant mentioning.

A difficulty was encountered in obtaining uniform fill of dry limas and pork and beans. In this case, a modification of the "lot plot" sampling procedure was used, first to determine the current operating conditions and, then, as a control procedure. It was felt that this could be used instead of $\bar{X}$ and R charts, "because of the intimate knowledge of the supervisory personnel with the problems of filling these products." It was found that, in some cases, filler speeds could be materially increased without altering the uniformity of fill. Also, fillers could be evaluated and the data and charts were used to confirm the need for new fillers. This particular type of data also yielded an estimate of the spillage on the filling lines.

SQC was applied by one canner to the sizing and canning of small whole onions. The problem lay in not getting the smaller onions out (the smaller onions are more valuable). After some equipment changes were made, $\bar{X}$ and R charts were used with some success. One interesting comment in the early stages of the data taking read as follows: "Our inspector fell and broke her arm-- no charts for two weeks."

REFERENCES

1. Bartlett, R. P., "Statistical Procedures Applied to Processed Fruits and Vegetables", National Convention Transactions, 1959.

2. Bremer, V. C., "Maintaining Net Weights Through Quality Control", Packaging Engineering, Vol. 2, No. 5, May 1957.

3. Donohue, F. S., "A statistical Quality Control Program for Packaging Lines", Special Report No. 15, American Management Association.

4. Donohue, F. X., "Statistical Application in the Control of High Speed Filling Machines", ASQC National Convention Transactions, 1957.

5. Filice, B. A., "Quality Control Techniques Available to the Average Canner", Proceedings of the 53rd Annual Convention, National Canners Association, 1960.

6. Gieseker, L. F., and Strasburger, L. V., "A Modified Lot Plot Sampling Procedure for Controlling Container Fill", ASQC National Convention Transactions, 1955.

7. Harrison, S., "A Packager Looks at Statistical Quality Control", Proceedings of the 16th Annual Forum of the Packaging Institute, 1954.

8. Hosking, F. J., "Quality Control Techniques Used in the Food Industry" ASQC National Convention Transactions, 1955.

9. Hudson, Wm. V., "Control of Quality Control", Talk given to the Canners League of California, January 8, 1960.

10. Kramer, A., "Food Processing Decisions by Multifactorial Analysis", ASQC National Convention Transactions, 1959.

11. Kramer, A., "Selecting an Efficient Sampling Procedure for Food Products", ASQC National Convention Transactions, 1956.

12. Kramer, A., "Quality Control Simplified for Practical Product Development", Food Engineering, December 1952.

13. Kramer, A. and Steinmetz, G., "Grading Raw Beans for Processing", Mimeo, University of Maryland.

14. Kramer, A. and Twigg, B. A., "Application of Statistical Quality Control Procedures to the Canning of Green Beans", Food Technology, Vol. XI, No. 12, December 1957.

15. Kornetsky, A., "A Quality Control Program for the Processing of Sweet Corn", Ph. D. Thesis, University of Maryland, 1956.

16. Kornetsky, A. and Kramer, A., "Quality Control Program for the Processing of Sweet Corn", Food Technology, Vol. XI, No. 3, March, 1957.

17. Lang, K. F., "Glass Jar Quality Control Really Works at Heinz", Food Processing, August 1955.

18. Lang, K. F., "Problems and Pitfalls in Acceptance Sampling of Glass Containers", ASQC National Convention Transactions, 1957.

19. Lee, E. F., "A Statistical Quality Control Approach to Weights and Measures Problems", ASQC National Convention Transactions, 1956.

20. MacPhail, I., "Quality Protection with Statistical Quality Control", Proceedings of the 53rd Annual Convention, National Canners Association, 1960.

21. McBride, John P., "State Regulatory Control of Prepackaged Commodities", Special Report, No. 15, American Management Association.

22. NCA, "Filling Operations in Canning", Proceedings of the Technical Sessions, 47th Annual Convention of the National Canners Association, 1954.

23. Reynolds, J. E., "A Practical Approach to a Weight Control Program", Packaging Engineering, Vol. 3, Nos. 11 and 12; Vol. 4, No. 1, December 1958.

24. Ritchell, E. C., "Control of Packaging Quality Using Statistical Quality Control Methods", Institute of Food Technologist Annual Convention, 1959.

25. Rowe, Sumner C., "The Effect of the Federal Food, Drug, and Cosmetic Act on Filling Procedures", Special Report No. 15, American Management Association.

26. Stier, H. L., "Quality Control Practices and Problems in Food Processing", Conference Papers Presented at the Quality Control and the Consumer Conference", Rutgers University, 1957.

27. Stier, H. L., "Statistical Quality Control Workshop", Proceedings of the 53rd Annual Association, 1960.

28. Tarner, M., Hall, B. A., and McDonald, J. G., "A Statistical Quality Control Approach to the Selection of Flavor Panel Members", ASQC National Convention Transactions, 1959.

29. Tarner, M., and Schenck, A. M., "Use of Extreme Value Control Charts in the Packaging Industry", Food Technology XIII, No. 5, 1959.

30. Twigg, B. A., "Control of Fill-in Weights", The Canning Trade, October 6, 1958.

31. Way, C. B., "Fill Control in the Canning Industry, ASQC Convention Transactions, 1955.

32. Weimerskirch, R. I., and Way, C. B., "A Mutual Approach to Quality Control by Can Manufacturer and Food Processor", ASQC National Convention Transactions, 1956.

33. Weiss, Wm., "Statistical Considerations in Food Standards", NCA Statistical Control Conference, Washington D. C., 1956.

34. "A Laboratory Manual for the Canning Industry", National Canners Association, 1954.

35. ASQC, State University of Iowa Section, Quality Control Training Manual, February, 1953.

36. Freedman, R., and Manshin, J., St. Louis Society for Quality Control, "A Basic Training Manual on Statistical Quality Control", 1951.

RAISING PRODUCTIVITY OF INDIAN INDUSTRY THROUGH QUALITY CONTROL

William R. Pabst, Jr., Chief Statistician
Bureau of Naval Weapons, Navy Department;
Formerly Adviser to the Government of India
Under the United Nations Technical Assistance Program

India, with its new and developing industries and with its national plans to raise the standard of living, presents a most interesting workshop for the effective utilization of SQC methods. It will be shown that these SQC techniques, when applied in the light of their effectiveness for the enterprise as a whole and for entire industries, that is in a "macro" approach, can be one of the most useful techniques in helping to increase productivity in the underdeveloped economies. Previous estimates that the productivity of Indian industry might be increased by as much as 20 per cent through SQC are confirmed in the present project with the finding that productivity gains from 10 to 40 per cent are potentially available. This experience arises from a United Nations technical assistance program in quality control established in 1958-59 under the auspices of the Indian Statistical Institute.

Background

Statistical quality control has been known in India for a long time through the work of Professor P. C. Mahalanobis, Director of the Indian Statistical Institute, and his associates. In the late Forties, Professor Mahalanobis invited Walter Shewhart to India for a series of lectures on Quality Control. Subsequently the Indian Society for Quality Control (ISQC) was founded. In 1952-53, Ellis Ott led a team of four experts, including Paul Clifford, Anders Hald, and Mason Wescott, on a three months United Nations training assignment, giving four two-weeks SQC training courses in the principal cities. Upon the basis of the interest and enthusiasm developed, the Government through the Indian Statistical Institute established SQC Units[1] in Bombay, Calcutta, Bangalore and New Delhi, and later in Coimbatore and Baroda. These Units were established to service industry on a fee basis.

Present Situation:

Today quality control is being used to an increasing extent in many of the major industries throughout the country, in the production of shoes, sewing machines, fans, cotton and jute products, castings, telephones, textile machinery, railway cars, pharmaceuticals, cigarettes, plywood, biscuits, bicycles and radios, among many others. These recent applications of quality

control methods have much in common, for they start with the search for those problem areas where statistical methodology can give the greatest returns; usually with the evaluation of outgoing product quality, the source of unusual scrap, or differential levels of efficiency. From these strategic entry points, the minimum web of control spreads where it is needed through the production process, sometimes to the control of raw materials, the improvement of manual processing, the elimination of inefficient or defective machinery and equipment, to the imposition of standards for production and inspection, or to other areas where management must exercise remedial or corrective action. This is called a "Macro" approach to distinguish it from one emphasizing methodology. This Macro approach is based on the notion that these tools of quality control must be utilized by management, for management, in the economic interest of the company as a whole if they are to be utilized effectively. Quality control or statistical techniques alone do not provide all the answers to successful operation. They do, however, provide management with the communication network on the technical side necessary for wise administration. The quality techniques found most useful in the Indian industries do not attempt single-handedly to solve tough problems or to apply genius methods where others have failed. They simply provide the mechanism of focusing management attention on previously hidden obstacles, and of utilizing the usually competent existing technical staff where solution is necessary.

An Example:

This approach is exemplified by the experience of a large plant in Calcutta producing sewing machines primarily for the domestic market in India and also for export to Egypt and Great Britain. It employs over four thousand workers and has a very competant technical staff. Attempts to use some of the tools of quality control over a period of time had made relatively little progress because the control charts were not understood or disregarded in the machine shop and assembly lines, and because successes in specific problem areas did not automatically spread. Besides a sewing machine is assembled from more than five hundred parts and it was clearly impossible with limited staff to utilize control charts for every dimension of every one of these.

When a broader view of quality control methods was taken, the initial step showed that no measure of outgoing quality was easily available. The corps of skilled final inspectors had no definite criteria of performance; instead they varied in judgment very widely; for instance, over an entire month, one inspector rejected over forty per cent of the machines coming to him, whereas another inspector dealing with the same flow of product rejected less than 10 per cent. Thereupon standards were developed utilizing the long time experience of operators for establishing procedural check-off lists and fixed routines of inspection. Demonstration machines for sharpening the borderline between good and bad were put to daily use. Within a short time the inspector judgment was narrowed to a five per cent range. In the meantime systematic exploration of the major causes of real rejections led to noticeable improvements in quality. One trouble, for instance, the undesirable or excessive "fork sound" was reported in some five per cent of the machines. By sampling the incomplete machines just after the installation of the fork, it was found that two of the assembly lines produced uniformly good product in this respect whereas the other six assembly lines were either inferior or variable. Since the same piece parts were used, it was relatively easy to study the assembly and adjustment process and ro rectify the conditions. Elsewhere in the sub-assembly department producing shuttles, it was found that one of the lines had three times the rejections of the other two, a situation previously hidden by the pooling of product despite 100% inspection. Again corrective action could be taken. In this way concern for

a dozen parts out of the total of five hundred made for striking increases in productivity, at the same time laying the foundation for a strong quality control framework and the effective utilization of control chart and other techniques where they were necessary. Management reported productivity gains of more than 10 per cent.

Explanation:

This growing national role of quality control in aiding industry confirms the expectations of Professor Mahalanobis that these methods first propounded by Walter Shewhart could contribute significantly in the national drive to increase productivity. He has been anxious to get as much technical help and assistance in quality control from other countries as possible, this making India unique as the only country in the world requesting assistance in this field through the United Nations.

As Adviser to the Government of India on quality control during 1958, I was assigned to work under Professor Mahalanobis the Indian Statistical Institute in Calcutta. My assignment was clearly to stimulate the national development of quality control in whatever manner possible, whether through the agencies of the Institute or through outside agencies. I soon found that the SQC Units of the Institute administered by Mr. S. C. Sen, Joint Secretary of the Institute was the really active force through which national growth in quality control might take place. Without the enthusiastic assistance and cooperation of Mr. Sen and the thirty or more statisticians and engineers in the Units, not much could have been accomplished. They made it possible to harness the growing enthusiasm of management and eagerly assumed ever-increasing new responsibilities in the field.

In the following pages, it is planned to discuss some of the factors allowing the changing concept of quality control to take hold. These factors include training courses , plant visits, industry-wide approaches, consumer goods evaluations and export controls. The viewpoint is, of course, that of an outside 'expert' looking to the means of securing a better utilization of the available quality control resources for a truly national SQC effort.

Methods and Techniques:

The six SQC Units of the Indian Statistical Institute provided a starting for the extension of quality control practices, inasmuch as the staff of some 30 statisticians and engineers had considerable plant experience and facilities for holding elementary training courses. Other starting points were explored. Advanced training courses were tried in an attempt to build a corps of industrial leaders, but this was disappointing. The Indian Society for Quality Control, limited mostly to Calcutta, and the Bangalore Quality Control Society in South India, were amicable but too limited in scope and energy to lead the way to vitalization of the quality control effort. The Indian Standards Institute supported quality control in its standards work and encouraged its members to join the SQC Units, and this was a helpful supplement. The National Productivity Council, coming into operation during the year, also encouraged the application of SQC through its publications and literature but was too new to exert leadership. The active methods of disseminating quality control principles was thus primarily through the training courses, of the SQC units,plant visits, and publicity developing from these.

Training Courses:

Two week training courses had been used by the SQC Units primarily on the model of the early courses given by Professor Ott and his associates. These courses had become heavily bookish and theoretical, failing to bring into focus the live and recent SQC experience developed in India through the work of the Units. A considerable amount of effort went into reorienting these courses bringing into them live current experience. Another important change was separating the two instruction weeks, and inserting a three-week interval between them. This split-week principle worked well indeed, for it provided an opportunity for the instruction staff to visit the trainees on their jobs and to suggest the actual means by which SQC could be implemented in the various companies. The second week of class concentrated on the actual progress made by the trainees in applying SQC in their plants. The visits to the plants, and instructional suggestions, generally flowed through plant management, providing the start of communications between management and the quality control groups. When management itself did attend courses, and they came in increasing numbers, the effect of training on SQC applications in the plants became much quicker and more certain. On no occasion did a manager attend a course without starting a useful application in his plant.

An Advanced Training course was given in night sessions in Calcutta over a three month period, using Dixon and Massey as a primary text. The results were disappointing for the more advanced statistical treatments did not help the industrial trainees to meet their own plant problems, which were primarily ones of position, organizational frustration and management non-response rather than statistical. Upon the basis of this experience it was considered that training at this advanced level was secondary to the other more active means of developing of a larger body of practioners.

Plant Visits and Follow-up

The most important avenue to the wider development of quality control was found in actual plant visits. These visits included those companies already members of the SQC Units, those whose students attended the training courses, and those otherwise interested. Over a hundred different plants were visited, some of them many times over. Each new plant visit was always in company with Unit personnel and was usually the occasion of some ceremony. But no matter how astute the observations during the visit, there seemed to be no lasting or tangible results without a fairly definitive follow-up letter explaining in some detail the steps that should be taken to initiate or improve the quality control program. Copies of a few of these letters are included in the Appendix.

These letters provided a communication channel between management to whom they were addressed and the quality control groups in the plant and in the SQC Units. They are of course very general but they did provide a tangible basis upon which implementing management directives could be issued. In well over half the cases, the suggestions were carried into effect, in many cases with good results. Copies of the letters were filed in each of the six SQC Units to serve as guide for use in similiar circumstances. Similiar letters have since been prepared by the Unit personnel as a preliminary step.

Typical Applications of SQC in a Plant:

These visits and follow-up letters are made somewhat easy because they are typical opportunities for the effective employment of quality control methods in most plants. Some examples are described under the following headings:

1. Balancing Production Efficiency: Jute weavers range in efficiency from 40% to 80%; spinning frames working on the same material produce surprisingly different amounts of yarn; foundry groups working with common materials produce only two-thirds of the good product of parallel groups; one sewing machine sub-assembly line has three times the defect rate as two parallel lines. Quality control methods show these differences in quantitative terms, analyze the non-common factors, and raise all units to a common level of efficiency, usually higher than the best before.

2. Reducing Scrap and Rework: In an automatic glass plant, scrap was reduced from a level of 15% toward the level of 1 to 2% that statistical studies showed to be attainable; foundry scrap has been reduced to a fourth; pottery and rubber tire seconds have been halved. Quality control starts with a cause-effect analysis of scrap, determines the most likely causes and seeks their elimination.

3. Standardizing Quality Judgment: In sewing machine evaluation, some checkers rejected four times as many machines as others, causing both poor outgoing product and unnecessary rework; similar differences in inspection standards were spotted in shoes, potteries, aluminum and in other products. Quality control helps to establish uniform standards and to control inspection accordingly.

4. Improving Product Quality: The outgoing quality of cotton textiles, jute, pharmaceuticals, lamps and other products improved without increase in costs, primarily through pre-engineering consumer requirements and controlling production accordingly.

5. Saving Raw Materials and Supplies: A pharmaceutical plant reduced the cost of fuel; a textile machinery plant found means to eliminate expensive imported materials while maintaining desired quality; jute mills have found means for utilizing less expensive grades of raw materials at the same time increasing outgoing product quality; all through the systematic observation, measurement and analysis. To these typical instances might be added those making maintenance procedures more effective, reducing waste, controlling weights, and a variety of others. For purposes of convenience these typical applications can be organized in the form of scoring sheet or check off list so that the places of maximum potential gain can be quickly noted during a plant visit.

An Industry-Wide Approach:

After visits to two or three jute mills, the problems disclosed and the suggested pattern of approach to quality control appeared to be much the same for all. Most mills faced essentially the same problems. It was therefore possible to develop a general approach on the basis of a few visits that would be useful and meaningful to all the mills in the industry. These approaches were called industry briefs and were circulated through the trade associations. They provided concrete modus operandi appropriate to Indian conditions as given in the follow-up letters, but of course lacked the personal touch and direct contact with top management. Nevertheless these industry-wide prescriptions for quality control attracted a great deal of interest and encouraged many of top management of the jute industry to attend training courses in quality control.

Similar industry-wide approaches were projected in other industries, plywood, pharmaceuticals, foundries and potteries. They provide not only for specific applications in industries, but also for the background knowledge that Unit quality control practioners need in approaching new plants and in standardizing their techniques.

Consumer Goods Testing:

The comparative testing of consumer goods purchased from the market place provided another avenue for exploring where quality methods might be advantageously used. This avenue was explored very slightly for in India there is yet nothing comparable to Consumers Reports [2] providing objective quality comparisons of consumer products on the USA market. Two projects were started, one dealing with ordinary glass tumblers and the other with box safety matches.

Common glass tumblers of the same price class, made by some eight different manufacturers were purchased from the market in Calcutta. The object was to score the quality of the different manufacturers and to use relative position in the quality scale as leverage with which to approach the manufacturers concerning the introduction of better controls in their manufacture. Most glass plants in India are small and use either hand or semi-automatic equipment. A direct approach to each plant would be time consuming and costly, for it takes just as long to convince a manager of a small plant on the need for better methodology as it does to convince the manager of a large one. It was reasoned that an approach supported by incontestable market facts might be quite effective. There would of course be additional leverage if brand name quality information could be effectively presented to the consumers as now happens in most of the Western countries.

The glass tumbler scores based upon a carefully designed sampling plan showed startling differences in quality as between the manufacturers, the objective score of the poorer being only about a third of the best products. The information was presented at the Indian Standards Institute meeting in New Delhi in coded values and also sent to each of the manufacturers. Time did not allow a complete follow through nor a real test of the effectiveness of this approach.

One side effect of the consumer oriented approach, however, was the development of a satisfactory test method. When the first token sample of glasses was examined, the examiners differed more in their judgments than the products. In the several successive rewrites of the instructions for evaluation, the differences between examiners were studied and narrowed down until the point when a outside group of examiners could attain essentially the same scores on the different glasses. This training in the proper preparation of a classification of defects or inspection instruction, based upon reducing the error of observation to satisfactory levels, provided invaluable training for actual quality work in the plants.

Safety matches were also evaluated from a consumer quality standpoint. This developed from the Unit training courses at first, for the small boxes provided an easily used class room device for preparing an inspection instruction and for evaluating the product of the different manufacturers. The quality characteristics examined were the number of matches per box, the number of broken or defectives, and the condition of the box and slides. The quality differences were so pronounced that the students could easily detect the areas of processing among the different manufacturers where better control of processing could be effectively employed. This information did get to the match manufacturers and there were persistent calls for assistance in quality control from them.

Despite their limited usage here, consumer goods quality evaluations provide an excellent starting point for stimulating the development of a national quality control program.

Export Quality Control:

There is some interest in India in the use of quality standards for the control of exported handicrafts through the Handicrafts Board, but most of the present handicraft products flowing into the export market depend so much upon artistic skill and workmanship, that not much more has been done than to assure the authenticity of the materials as represented; ivory rather than bone, and so forth. There appears to be little parallel to the Japanese case. In Japan exports play so important a part in the domestic economy that rigid quality control of exports in the post-war market has been a most effective instrument in extending the internal use of quality control techniques.

In India one unfortunate case of export quality control was observed in the case of druggets, the hand-woven rugs made of tanners wool. At the same time attempts were made to impose export quality standards, the increasing cost of tanners wool and the development of synthetic fibers abroad for rugs in this price class almost completely destroyed the foreign market. Quality control has not excaped from criticism in this fiasco.

National Meeting on Quality Control:

The National Conference on Quality Control held in Calcutta at the end of 1950 exemplifies another method of developing and extending SQC in India. Some 250 people, representing top and middle management and SQC practicioners, attended the three day session. Over 60 papers and reports were given, mostly dealing with the development of recent quality control programs. The papers were drawn from over 20 industries and reflected the considerable progress made in the application of quality techniques. The conference was a most enthusiastic one, and the spirit of accomplishment on the part of those who participated was dominant. The conference was jointly sponsored by the Indian Statistic Institute and the National Productivity Council.

Conference of this kind accomplish many purposes. First they provide illustrative material which others might follow. Second, they bring quality control work of individuals or groups within a plant to the attention of their own management in a manner not possible in internal reports. Third, they engender confidence on the part of those who have done the work. And last, they engender group morale on the part of all those engaged in quality applications arising from the scope and intensity of the applications. All these are necessary in the development of a national program.

The particular conference in December 1950 provided a great deal of factual material on the quality control developments in India. This together with other observations are summarized in the following notes on some of the industries:

Sewing Machines: A plant-wide application has helped to save 10 to 20 per cent rework, increased efficiency of processes by balancing assembly and sub-assembly lines, and helped to raise productivity by more than 10 per cent.

Biscuits: Provided means for control of biscuit and packaging weights and for improvement of products.

Soaps: Disclosed existence of nearly fifty per cent process run around owing to faulty cake solidification and cutting. Control charts used extensively in laboratory tests. Methods helped to improve packaging quality.

Cement: Revealed that over 75 per cent of bags were shipped under-weight unintentionally as a result of automatic weighing machine errors and faulty sampling controls. Methods helped to reduce this percentage to less than six, while raising the average bag weight by less than two pounds. Extensive applications considered in control of particle size in slurry and in the balancing of fuel consumption of kilns.

Textile Machinery: Statistically designed experiments led to reduction of blade breakage and conservation of expensive imported materials. Many partial SQC applications in some of the plants.

Telephones: Large numbers of control charts in machines shops did not keep defective units from assembly lines. A more dynamic quality approach starting from evidences of unsatisfactory product, eg. noisy dials, seems to be working.

Typewriters: Large numbers of machine shop control charts but no effective controls in assembly operations and in final evaluation where needed.

Abrasives: Effective plant-wide quality program developed with reduction of wheel breakage and paper scrap, replacing the spotty and partial control chart operation previously used.

Prospect for Future Development:

Letters and reports received from India during the past year indicate that the promising start in the use of SQC methods has continued, although not without some backward slippage. In the progressive northwestern area, the Baroda Unit has all the plants in its region as members, while the Bombay Unit has a waiting list. In the relatively unindustrialized north center, the New Delhi Unit reports heartening progress with new returns in the sugar industry and into hotel applications. In the South, the Bangalore Unit reports the addition of many new member companies. The Calcutta Unit, however, reports some setback with respect to its membership possibly attributable to the absence of the unit head on training in England. Altogether there appears to be a lively and continuing interest in actual applicatinns.

This interest still does not appear to be reflected in the size of the membership roles of the quality control societies. One may have to accept the fact, however reluctantly, that the ASQC pattern may not be the only way of developing a national quality movement.

In this last year, too, two of the unit heads have had six months on the job training in England and others are expected soon in this country. There is every hope that they will return to their homeland with renewed enthusiasm and with clear examples of effective SQC applications to follow. Their difficulty besides that of finding good single plant examples is that in this country and in England there is no comparable governmental group vested with national responsiblity for SQC.

Jute Products: Industry-wide applications appear to be in prospect aimed at raised spinning and weaving efficiency (in order of 10 per cent) reducing raw material costs, and raising outgoing product quality. Some twenty mills out of the national total of 100 have active SQC programs.

Cotton Textiles: Many applications throughout the 500 mills in the country are spurred by the three Textile Research Associations, by private consultants, and by the SQC Units. Some mills report 20 per cent increases in weaving efficiency through quality control, reduction in waste, and product improvement.

Foundries: Many applications in large and small foundries thoughout the country, many of them reporting very significant reductions in scrap and defective castings, and increases in production. One large foundry reported reductions in defective castings from 30% to less than 10% through reinforced molds, improved cores, and better controls of sand initiated through the quality efforts.

Bicycles and Accessories: There are a strong SQC applications in several of the plants. One reports reduced painting rejections from 25 per cent to less than five; another reduced rim rejections from 35 per cent to less than seven.

Glass: In a large automatic glass plant, SQC helped to reduce the rate of rejections from 16 per cent to less than half, with great increases in the quality of product especially that going to the pharmaceutical industry.

Pharmaceuticals: Many applications in a large number of relatively small plants. Significant improvements have been reported in product quality especially in the elimination of foreign matter in vials, correction of loading problems of bottles and caps, and the provision of uniform weights for tablets and capsules. Increased emphasis on quality standards in related bottling sidelines.

Potteries: In a large pottery, very significant reductions made in the amount of scrap or broken plates and the proportion of seconds through control charts and other statistical devices. Differences in roughness of automatic ovens revealed as cause in addition to lack of training of moulders.

Shoes: In a large factory SQC methods detected and helped to eliminate the 10 per cent substandard shoes previously sent to the market.

Aluminum: In a large rolling mill, quality methods helped to reduce final rejects and scrap through improvements of tempering methods and rolling processes, adding a dynamic quality note in a plant where control charts had been previously used primarily for plate thicknesses. Quality committee covering all departments established to coordinate design as well as production improvements.

Electric Lamps: In a relatively small plant statistical design of experiments led to greatly improved length of life of bulbs, and reduction of scrap of automatic bulb making machines.

Cigarettes: Quality methods led to effective control of machines producing imperfect cigarettes; controlled moisture and weights. Methods used in many plants.

Letter to the Manager of a Fan Factory

My preliminary observations, based on the brief visit and similar situations elsewhere, are that the first important step of the quality control person should be to systematize and utilize the final inspection observations, then systematize incoming inspection, and finally the machine shop. Some details on these steps are given in the following: -

1. Final Inspection: Situation: at the present time there is a considerable percentage rejection of fans at the final inspection. Speed in RPM and power consumption are checked. Causes of rejection of individual fans are noted in the inspection book, but no summary is available. Many of the table fans were apparently rejected for poor regulation, the cause of which should be easily isolated and correctable. Apparently the different lines vary considerably in the rate of rejection.

Action suggested: Final inspection records should be summarized over the past month for each line as a guide to action. The rate of rejects for each line should be plotted on a control chart. The reasons for the individual rejections should be summarized and investigatory action should be taken with the most important ones. This may show where manufacturing, processing or assembly quality controls are necessary.

The parallel was encountered in sewing machines. Enumeration of causes of rejection after inspection was standardized led to the tracking of reasons for shuttle and fork sounds. Corrective action is being taken in shuttle sub-assembly for the first, and on the assembly line for the second. These strategic points were located after considerable study.

Additional attention might be given to the inspection procedures for fans: There was no indication that fan blades or guards were checked at the final stage or that possible 'minor faults' were noted. Once the final inspection is strengthened and stabilized, and the obvious causes of rejection corrected, a systematic measuring of outgoing quality should be established by mean of a small sample check inspection just before packing. This will provide a real control on production test and assembly operations.

2. Incoming Material: Situation: Inspection of incoming material is apparently conducted, but there is little indication that records are maintained or utilised. Control of incoming material can become one of the most effective and important controls in the plant. Some material now is sorted out and returned to the supplier (fan blade guards), but the quantity of this return, the reasons for defective classification, or the usability of the remaining stock is not assuredly known. Aluminum sheets were said to be returned to the supplier for deep scratches, but no indication that actual returns had been made was available. Ball bearings are checked apparently on a sample basis, but explicit sampling plans or the basis for lot acceptance and rejection were not found.

Action suggested: For the important incoming assemblies and raw materials, a systematic effort to evaluate incoming quality for each supplier should be made with the assistance of MILL STD 105a sampling tables. For each such product a careful inspection procedure will need to be established (sometimes in consultation with the supplier) and a lot-by lot check inspection made and recorded. Information thus obtained should enable the determination of practicable acceptable quality levels for each type of material.

Such control of the quality of incoming materials will (1) make it possible to isolate non-acceptable suppliers (2) return reliable information on the causes of rejects to suppliers to secure corrective action in the suppliers' plant, and (3) reduce processing operations due to substandard materials (e.g. scratched fan blades).

3. Machine Shop. Situation: Considerable screening or one hundred per cent inspection exists after many machining operations. The present record books are difficult to interpret, but they seem to show (1) that rejection is usually small and confirm operator's statements that it is about one per cent, and (2) there is considerable difference among operators on the same machines with respect to scrap and rework rates.

Action suggested: Attention might be devoted to this after a satisfactory system is operating on final inspection and incoming material. Important thing first is to find from final inspection where processes need closer control than now provided and concentrate on these strategic situations. Thereafter a control chart inspection on the machine, or a sample lot inspection check, can be substituted in turn for most of the 100 per cent inspections now practised. The object of the sample checks is to control the process and prevent the production of defects, rather than finding them by detailing and later correcting the observed defects as at present.

These broad suggestions are offered as a general guide for the development of a quality control system. As you know, progress in the installation of a quality control system may be slow, but it must be determined. The object must be to save in costs, better utilize facilities, eliminate unnecessary plant run-around and rework. Gradual accomplishment of the above three steps should work toward the accomplishment of the cost-saving objectives and lead to a complete quality system.

Sincerely yours,
Sd: William R Pabst, Jr

Letter to a Jute Mill Manager

Our visit to the mill yesterday was a very pleasant experience. It was good to see your keen interest in these new methods of quality control and to find how neatly these statistical methods dovetailed with your past measures of control. Under your progressive guidance, we look forward to very useful applications in your mill.

As we discussed in our conversation, quality control methods cannot be imposed by an outside authority, but rather must develop within a plant. These methods put in scientific form much of the information that has been ordinarily collected in the past. These scientific methods attempt to search out the strategic factors that dominate a process and to develop methods of control of these factors. The extent of control that is necessary depends upon economic considerations, for these methods must prove themselves in increasing productivity, saving costs and improving the quality of output. Our plan of implementation is as follows: -

1) As agreed in our meeting, the first application of quality control methods will be in the spinning stage of the process. The overall purpose here is to utilise the present spinning equipment as effectively as possible in turning out yarn, to indicate the men, frames and spindles that are significantly below the rest, and to study maintenance procedures on the frames with a view to their optimisation.

The first task in working toward this objective is to evaluate the present state of efficiency frame by frame. To do this, a systematic snap check reading of ends down will be taken four times a day at random intervals. This will be similar to but much more extensive than the initial reading we took yesterday in which we counted 23 ends down on one frame and 2 ends down on the frame next to it working on the same yarn. This information recorded over a week's time will serve to determine the basic level (in ends down) at which frames might be expected to operate for each class of yarn. The frames departing from this determined level either continuously or spasmodically will then be indicated for special attention by the spinning master and maintenance crew.

The second task will be to make frequency distributions of end breaks spindle-by-spindle for the frames lowest in efficiency. The object here will be to designate those spindles giving performance very much inferior to the others and to secure their correction. It will also be to study correction needed for these poorer spindles in the light of a preventive maintenance policy.

Operations connected with spinning may also be studied under this first application in connection with the tracing of the reasons for departures from established standards of performance back to their basic causes, and to necessary corrective action. This first application should be reasonably well established at the end of one month.

2) The second application is planned for the evaluation of outgoing material and will be initiated during the second month. The objective here will be to evaluate product quality to ensure that the entire process of mixing, spinning, weaving and inspection results in the quality of material that is intended. If not, the

information will be at hand to determine the strategic processes and apply whatever measures are necessary to control these.

The mechanism of this second application will be to establish a check evaluation based on a sample of 16 to 32 cuts per day. These cuts will be examined on a systematic scoring scheme whose detail will disclose wherein process troubles may exist. Actual procedures will be spelled out in due course for implementation at the mill.

3) The third step to be taken in the third month will depend to some extent on the findings of the second, but will probably be a systematic application of quality control at the weaving stage. For this measures of the efficiency of looms on the different fabrics will be studied to determine reasonable standards, and to note significant departures from these. A concerted attempt will be made to narrow the differences in efficiency among the individual weavers so that the looms may be utilised closer to their capacity.

Data already available from efficiency payments to the individual weavers may be utilised in this aspect of the work, and will be supplemented by snap check readings as necessary. More systematic control of faults at the looms may also be found necessary.

4) The fourth and last step interchangeable with the third will be to introduce more systematic control of batching based on available information on the relationship of spinning and weaving efficiency and batch cost. The purpose will, of course, be to work continuously toward the optimum batching.

The mechanism will be to determine the stabilised spinning and weaving efficiencies related to the different batch compositions and their costs, and to study this relationship through correlation techniques. This information should supplement the experience and skill of the mill in arriving at optimum batching.

This four month program is designed to keep the quality control application moving along at a consistent and reasonable speed and to allow for the training of mill personnel during its course. Within a year's time, it is considered that the mill should be able to go forward on its own with the application of quality control methods. The primary endeavour of the SQC Unit assistance will be to put the mill in such an independent position with respect to quality control.

It is hoped that this programme will meet with your favour and that you will guide it to an economic and practicable.

Sincerely yours,

Sd: W R Pabst, Jr.

Letter to a Soap Manufacturer

The lively discussion of the use of Quality Control techniques which may contribute to your developing programme was most interesting to me. You and your assistants are apparently aware of the great gains that can be made through SQC in the several areas. For purposes of record let me recall them as follows:

1) Elimination of unnecessary scrap at the cutting of the soap cakes:

As indicated there are nearly 30% turn-around at this point, thus involving unnecessary rework and diminution of your production capacity. What is necessary here is a systematic record of the amount of rejects at each cutting bench, the differences between time as measured in days or weeks, and the difference between operators. This information may be gathered on a complete basis or by sampling, taking half hour period during the morning and afternoon shifts to provide the necessary statistical information. The data should also include classification by cause of rejection and this should lead you soon to corrective action needed. My impression based upon observations in other plants as well as of your excellent staff is that you will be able to reduce the amount of rejection at this point from the present 30 to less than 10%.

2) Difference in inspection standards:

As you remember, we watched the different inspectors rejecting the soap cakes after stamping. It appeared that the standards of these inspectors differed very much. Presumably some of the material rejected by "Tough" inspectors is passed by others; consequently some poor material goes on to wrapping. Quality Control should introduce proper standards for inspectors on all the lines so as to achieve a uniformity of acceptable product.

3) Wrapping of soap:

Quality Control can be usefully applied to the wrapping of soap in order to achieve uniformity among the different wrapping machines throughout the plant. It was claimed that there was some 5 - 7% loss at wrapping, possibly higher in the case of small soap cakes. Some of this loss may be attributable to different machines and to different maintenance. In other similar cases we have found that the difference between individual wrapping machines account for a substantial part of this loss, and that this generally can be corrected.

4) Outgoing Quality of Product for Soap:

The measurement of outgoing soap quality including the packaging of soap cakes that now get through 100% detailed inspections might easily be measured. The finished cakes are now put into boxes for shipment; a valid procedure would be to take a randum sample of about 32 cakes per shift for each machine. These cakes should be examined carefully and scored for defects in wrapping and defects in the cake itself. My impression was there is a percentage of both defective wrapping and dirty or spotty cakes now going out of the plant in most kinds of soap. My impression also is that most of these cases are unnecessary and can be avoided.

5) Weight of Soap Chips:

Quality Control study should be made of the weight of soap cakes and their packing material. Your present procedure seems to be biased on the heavy side for the girls tend to fill up a given weight. It is possible that the packages average nearly a quarter ounce overweight. This can be easily checked through quality control techniques and procedures set up for arriving at very close weight distribution about the average that you desire.

As to the other places for quality control as in the chemical control of soaps, it appears to me that this is now well in hand through your laboratory. If the more obvious things are done in the wrapping and cutting plant it may be that Quality Control techniques may be used in Laboratory in gaining long run as well as batch to batch control of the basic materials.

As for the raw materials sampling, special problems are encountered here especially in your need to sample each barrel. Probably better control here is not of our immediate concern but I suspect that you will do very well to add tests of four individual barrels from each shipment in addition to the test of the composits that you are now making. This test on individual barrels will give you the measurement of the differences among barrels; better information of the character of the lot thus obtained may help in many ways in the batch control of your processing.

Sincerely yours,

/S/ William R. Pabst, Jr.

Letter to the Manager of a Pottery

Your plant has many modern features with its beautiful oven and the automatic controls. Your products especially on the sanitary-ware side are in great demand. You are handicapped to some extent by the raw materials available, by the amount of scrap and rejects that occur, and by the capacity of oven space. Quality control methods may help to overcome some extent each of these limitations.

1. Raw materials: The iron content of the clay that you receive from local sources comes as the result of faulty processing of the clay. The iron tends to discolor the china after baking. Since you apparently cannot reject lots of clay, the only thing that you can do is furnish your supplier with an account of the impurities of the clay. At least two samples from the incoming material might be sent to a laboratory for analysis, possibly at one of the colleges or governmental labs in your region. This will establish whether your supplier is consistently faulty, or has good and poor shipments. If the latter, you will do well to try rejection of the poorer lots received. If the supplier is consistent, then bring to his attention the condition of the material received may have the effect of getting him to improve his process in this respect. Ultimately he will have to improve or you will have to install processing machinery if satisfactory raw materials are to be obtained. Apparently the quartz and felspar are satisfactory.

2. Scrap and rejects: The high rate of rejection on both the sanitary-ware and on the china need thorough study of cause-effect relations. Many of the troubles apparently can be traced back to the molds or to the molders, but a method of tracing this information systematically is needed. The first thing that is necessary is a scoring system or check-off list of the observed kinds of defects so that each days production can be scored by the number of good pieces, and the number of defectives of various categories. The second thing that is needed is a method of identifying the original mold and molders (or group of molders) so that the comparative scores of these groups day by day can be obtained. This segregation is of utmost importance, for if experience in other potteries is relevant, it may be found that some groups do an excellent job whereas others do an imperfect one. The object is to get all groups at the same level through training, instruction or other procedures, but this cannot be done until it is learned which groups need such improvement. Additionally the different rate of defects for the different kinds of china may provide the clue as to necessary changes in processing.

Incidentally the present method of sorting into firsts, seconds and rejects, needs careful review to remove what appears to be large differences in inspection judgment. Visual samples of just good and just bad pieces of each kind of usual defect may be helpful in bringing all the inspectors into a common judgment. Assembly of material to be sorted into discrete batches may also be helpful in avoiding this commonly found bias.

3. Utilization of the oven: Since the space in the oven is the critical and limitating factor in production, your efforts to utilize as fully as possible the available 32 cubic feet of each truck should be continued.

You might measure the actual cube used by each of the kind of loads you now make and start exploring as to how to raise the lowest of these. The short-time advice of a geometrician might be most helpful to you in this regard.

My feeling based upon the short visit to your plant is that very considerable progress can be made especially in reducing drastically the rate of defects.

Sincerely yours,

/S/ William R. Pabst, Jr.

Summary and Observations:

1. Statistical quality control techniques properly applied can be one of the most effective means of increasing the productivity of established Indian industries.

2. Individual plants in the increasing variety of industries are starting to use these methods.

3. Management of Indian industry is more responsive to the application of these techniques, provided they can be shown in concrete terms where they apply to their own enterprises and to Indian conditions.

4. The SQC Units of the ISI, with their training and follow-up facilities, in-plant assistance and promotional activities, provide the strongest nucleus from which a national quality control movement might spring. The USA pattern of SQC growth through a strong professional society does not appear to be easily transplanted to Indian soil, nor does the Japanese pattern of development through strong controls of export quality.

5. India provides a most suitable training ground for those interested in the "macro" application of quality techniques. In addition it has great resources of mathematical and statistical skills which can be turned to industrial pursuits and which might serve broader purposes in the free world.

6. The possible use of quality control techniques as one of the most effective instruments of technical assistance to the underdeveloped nations of the world should be of strong and increasing interest to the members of this society. 3

Bibliography:

1. Ellis R. Ott, "Quality Control in India", United Nations Report, 1958.

2. Consumers Reports, published monthly by Consumers Union, Inc., Mount Vernon, N.Y.

3. William R. Pabst, Jr., "Use of Statistical Quality Control in the Industry of Underdeveloped countries", Industrialization and Productivity, February 1960, United Nations, N.Y.

CALIBRATION REQUIREMENTS FOR GOVERNMENT CONTRACTORS

William B. Rice, Head
Measurement Standards Department
Aerojet-General Corp., Azusa, Calif.

Mr. Chairman, Fellow Members and Guests of the American Society for Quality Control:

It is a privilege and a pleasure to address this distinguished audience, since my subject, "Calibration Requirements for Government Contractors," reminds me of the rabbi and the priest who met on the street one day. These two gentlemen of the cloth were old friends, often engaging in spirited theological discussions. On this particular morning the discussion became especially heated until finally one of them said,

"Although there are many subjects we can not agree upon, there is one I am sure we can."

"What's that?" asked the other.

"Well, you serve the Lord in your way..."

"Yes, yes," said the other.

"And I," concluded the first man, "serve Him in His way."

Looking through recent issues of Industrial Quality Control, I can find no serious discussion of calibration. This is why I feel that I am "serving the Lord in His way" when I bring it to the attention of this Conference. I feel very strongly that this subject should be a serious concern of the American Society for Quality Control, because, involving, as it does, the accuracy, precision, repeatability and compatibility of measurements, it obviously holds a central and imperative place in quality control and reliability.

I asked the Head of our Reliability Control Staff at Aerojet how much effect the accuracy of test measurements has on the reliability of a product. He said, "About 75%." I think it is at least 75%, because no matter how carefully your tests are devised, how scientific your sampling, how realistic your environmental conditions, unless your read-out instruments are telling the truth, what assurance do you have?

Somebody in the Pentagon apparently believes that the answer to this question, "What use are measurements if you don't know whether the measuring instruments are telling the truth?" is vital to national defense. In 1953 the Naval Bureau of Ordnance (now BuWeps), allocated substantial funds for establishing a calibration program for their own installations and their missile weapons systems contractors. This program has now established two primary standards laboratories, one on the East coast and one on the West coast, together with a dozen or more secondary laboratories in major areas across the country. These facilities are available to Navy contractors under the conditions imposed by the program. Aerojet-General has been one of the first and most active participants.

The Atomic Energy Commission has long had a superb measurement standardization program with a magnificent primary standards facility at the Sandia Corporation in Albuquerque, New Mexico.

The Air Force has also become concerned. Last August an official opinion was issued by the Judge Advocate of the Air Force which says,

"MIL-Q-5923c (paragraph 3.3) requires the contractor to check gages and

other testing and measuring devices with suitable measuring equipment at established periods to assure continued accuracy.

"As MIL-Q-5923 is a contractual requirement, the contractor is completely responsible for maintaining an adequate reliable source of calibration. The National Bureau of Standards being the only reliable source of calibration in this country, the official opinion of the staff Judge Advocate is that the contractor is contractually obligated to maintain a source of calibration traceable to NBS. Traceable, as used here, is comparable to the tracing of an electrical circuit for continuity. The path may not be direct, but leads to the source in the shortest possible way.

"The new DOD Specification MIL-Q- 9858 superseded MIL-Q-5923, and although the phraseology of the two specifications differ, the rights and obligations of the contracting parties remain the same."

The Air Force has a primary standards laboratory at Wright Field and has established, or is establishing, secondary laboratories in each Air Materiel area. They do not, however, make these facilities available to Air Force contractors. It is the responsibility of each contractor to fulfill his contractual obligations in accordance with the policy just quoted.

What does the word "calibration" mean as applied to the contractual obligations of Government suppliers?

In the broad sense calibration is a procedure for assuring ourselves, through a chain of certified standards, that our measurements are as good as the instruments we use permit.

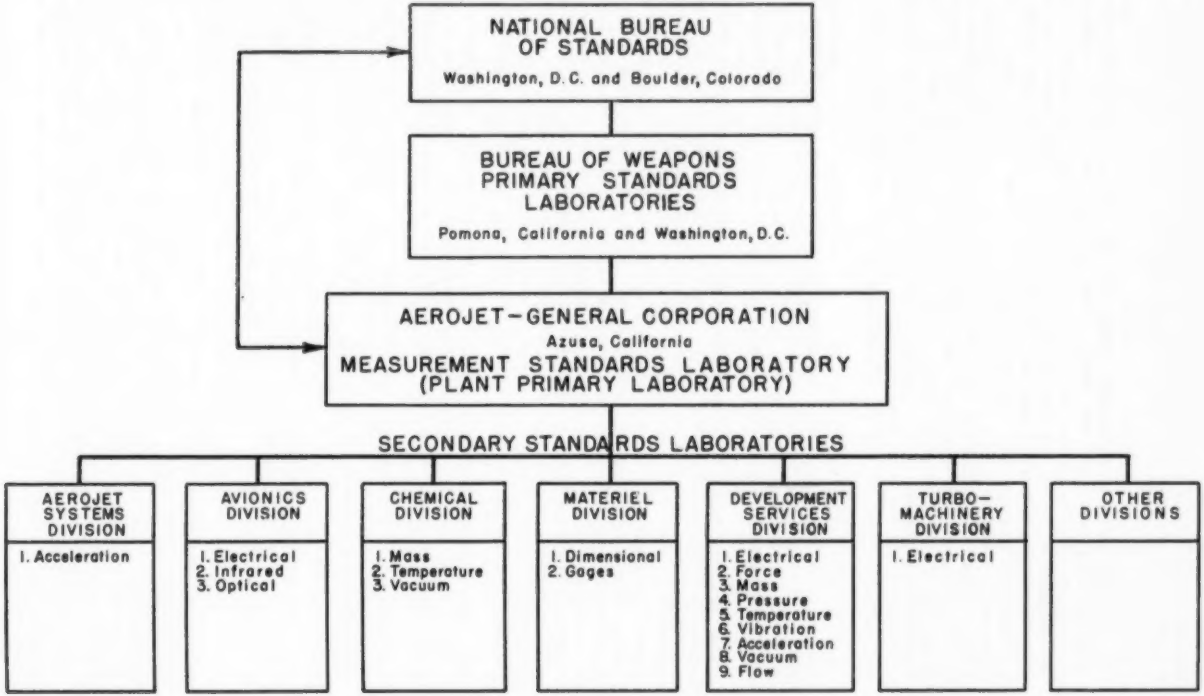
For instance, a micrometer, properly maintained and used, should be able to measure a dimension to within one-thousandth or one-thousand millionths of an inch. A supermike takes it down to one-tenth of a thousandth or one-hundred millionths. A set of Class "A" gage blocks goes to about ten-millionths and a gage block comparator to one-millionth of an inch. There is, however, a basic difference between a "mike" and a comparator. The mike will measure an absolute dimension, such as one or one-hundred thousandths of an inch. A comparator measures only the difference between the unknown thickness of a gage block and the known thickness of a standard block. Most high-order accuracies of measurement are made by comparison with a standard; hence, the value of the standard must be known with an even higher order of accuracy; that is, it must be certified through local standards which, in turn, are certified through a chain leading finally to the National Bureau of Standards.

How is this done?

At Aerojet we have gone about it in the way shown in Figure 1. At the top, of course, is the National Bureau of Standards, the legal custodian of United States National primary standards of measurement, and the final authority in transferring accuracies from the National primary standards to industrial and local standards. In the Aerojet, Azusa, plant, we send our standards (resistors, cells, potentiometers, gage blocks, weights, etc.) to the National Bureau for certification at regular intervals; or to the Western Primary Standards Laboratory of the BuOrd Program at Pomona, just seventeen miles away, whose standards are in turn certified by N.B.S.

Thus, in our Measurement Standards Laboratory, we maintain plant primary standards with a direct link to N.B.S. This is the finest laboratory at our Azusa plant, with two separate rooms - the general laboratory at $72^{\circ} \pm 1^{\circ}$ and the Metrology Laboratory at $68^{\circ} \pm 1^{\circ}$. Relative humidity is controlled between 40 and 50% with dust control to five microns/cm^2. The floor is separated from the rest of the building on a floating base; tables and benches are specially designed to minimize vibration. Equipment is the best that money can buy in the U.S., England, Switzerland and Germany, with high-accuracy capability to certify standards in the measurement areas of length, mass, time, temperature, pressure, vacuum, acceleration, vibration and electrical-electronic. By the end of this

SCHEMATIC OUTLINE OF MEASUREMENT STANDARDS PROGRAM



year, we expect to add force, flow, optics and infrared, thus covering practically the entire spectrum of measurement requirements.

We have also established plant secondary standards laboratories in various convenient locations in the plant, in special air conditioned areas where particular instruments are assigned as secondary standards to be used only for calibration purposes. These secondary standards are certified by the Measurement Standards Laboratory, and used to calibrate working, test and inspection instruments and equipment. Records are kept; periodic recall for calibration or certification is required; and stickers and reports are used to evidence calibration status on all equipment. Thus, the legal requirement upon Aerojet by its military customers is fulfilled. The chain is directly traceable to the National Bureau of Standards.

This program has cost a lot of money and effort. Naturally, we feel obligated to meet our legal commitments; but over and beyond this, what advantage do we gain from such a program? Top management at Aerojet would sensibly not spend the money unless it were profitable.

My next statement may startle you. It is common experience to find, in a plant where accuracy is not properly organized, that 30 to 50% of the measuring equipment used is not giving true answers. Let me cite a few extreme examples.

We found, for instance, that one of our vendors had surface plates which, when calibrated with an autocollimator, looked almost like wash boards. No wonder we could not get acceptable height or flatness measurements from them!

Last summer we had the Los Angeles County Bureau of Weights and Measures come to calibrate and seal our scales, of which we have 70 in the plant, a job which took them three days. During that time, I became well acquainted with them and they told me this story about one of the other plants they visited. A man there was making some precise weight measurements. At one time he had needed a two-pound weight; so he brazed two one-pound weights together, apparently not worrying that this made the two pound-weights weigh two pounds and three ounces. Of course, the Los Angeles Bureau confiscated this home-made weight!

I must admit that even occasionally Aerojet is at fault. We found a standard cell in one of our facilities which, after being checked daily for a month, had a graph looking like the Rocky Mountains instead of being flat and level.

I could name case after case of this sort from my own experience. We are happy to say that after a year's hard work on our program, less than 5% of our instruments are found to be out of calibration, that is, more than 95% of them give us readings with means and dispersions that are within the manufacturer's specifications.

A case history may clarify my meaning.

One of our projects calls for tight control of air pressure. During fabrication of these items they are pressure-tested on Heise gages in the Manufacturing Division. Now, these Heise gages are laboratory instruments with a nominal accuracy of + 0.1%. However, when used for production tests on a 16-hour a day schedule, they cannot be expected to maintain that accuracy. Our calibrations of these Heise gages reveal that they are 1/4% instruments under these conditions. Although not maintaining manufacturer's specified accuracy, they are perfectly acceptable for the use to which they are put. Furthermore, we know their accuracies; we know when they deviate from permissible limits, hence we know that the tests we make are good. These gages are therefore approved for these specific tests, but not for any others where higher accuracies may be required.

We can not make these decisions nor have this assurance unless we have a chain of certified accuracies leading up to the National Bureau of Standards.

As you can see, all this goes back to N.B.S. When producers of defense

material can trace their accuracies back to the National Bureau, I am sure we will find that our percentage of missile successes will improve substantially. This we badly need.

With our defense system, where many contractors cooperate to produce a variety of subsystems in a number of plants all over the country, it is essential that, for instance, the measurements made in Azusa be compatible with those made in Baltimore or Denver or Los Angeles by someone else. The chain of calibrations, leading back to the National Bureau, when carried out with skill and sincerity, should achieve this compatibility.

Thus, the National Bureau of Standards carries a heavy responsibility. Unfortunately, its present facilities do not seem to be adequate for the greatly expanded job that has to be done.

A few years ago the Bureau could go quietly about its work, maintaining the national length, mass, temperature, pressure and electrical standards; certifying clinical thermometers, gage blocks, standard cells and other more or less routine jobs to reasonable tolerances. Now, suddenly, because of Defense requirements for calibration, they are called upon to certify ten-fold the volume of equipment from anyone who requests it, at all levels of accuracy, and at fees that do not cover the cost. In addition, they are asked to certify things that were never standardized before. A survey of calibration facilities in the United States, recently completed by the Aerospace Industries Association, tells the story. For instance, N.B.S. has nothing yet (though development is under way) for certifying temperatures below 12° K and over $4,500^{\circ}$ K. Inadequate in shock and complex vibration. Gaps in micro-wave and radio frequency coverage. Incomplete in infrared and nuclear radiation.

Clearly, N.B.S. also has a large and urgent research job to do in the exotic areas that have opened up in recent years. And more are coming: ion drives, solar batteries, space radios, new metals and plastics, all subject to the strains and stresses of extreme velocities, temperatures and pressure. Furthermore, accuracy requirements in missiles are rapidly becoming impossible in the present state of our knowledge. For instance, at Aerojet we have right now component parts that carry machining tolerances tighter than gage-maker's tolerances: holes spaced within 2/10,000 of an inch of true position on a thirteen inch circumference; shaft diameters +0 -1/10,000; bores in line within one-half a thousandth in 24 inches, and so forth. As we progress to the need for still greater accuracies, the concept of a standard meter becomes obsolete; a more satisfactory basic standard is a wave-length of light. In this area of basic standards with ultimate accuracies, the National Bureau of Standards must spearhead research. All at once their job has become overwhelming.

Now, what do they have to do it with?

They have a staff of devoted men who have given their lives for this Cause - many of them already past retirement age. They have done wonders with poor housing, inadequate budgets and maddening interruptions. Cranberry poisoning scares divert their precious, limited facilities from the urgent National Defense job. Do you realize that the operating and construction expenditures by N.B.S. in 1958 were only $29,000,000? Just a fraction of what Uncle Sam pays for one ICBM program.

The Bureau is well aware of its responsibilities. I quote from their latest published report:

"The responsibility for the Nation's fundamental standards of measurement demands a continuous search for new scientific knowledge and for improved measurement processes. As new areas of science and technology become active or productive, new standards must be developed, adequate measurement methods established, new instrumentation devised, and appropriate calibration services provided. All this must be done while advancing the established areas of measurement and in such a manner as to maintain and insure continued consistency and coherence within the whole complex chain of measurement."

In view of the situation as it now exists, the National Bureau needs much wider public support, more funds from congress, and complete freedom to engage in specific research projects that are immediately related to their legal responsibilities. Anything that the American Society for Quality Control can do to help the National Bureau would represent a most constructive contribution to the National defense effort.

If the National Bureau is to be free to do its basic research job, getting out from under the load of routine calibrations, there must be provided facilities in each geographical area of the country for the calibration of standards equipment used by industrial plants. For instance, a really adequate commercial calibration laboratory in the Los Angeles area would enable companies like Aerojet, the aircraft manufacturers, and others to get assurance of measurement accuracy much more conveniently and more quickly than otherwise, as well as relieving the Bureau's burden.

The need for commercial calibration facilities is acute. By this I mean facilities that maintain certified standards and, with appropriate environmental controls, compare, test, and inspect equipment against these standards. There are a number of excellent commercial laboratories in the Los Angeles area, but few of them are set up to perform this particular crucial task. I have personally surveyed many of them, and the Air Force has done a thorough survey. The results have been disappointing. I have seen millions of dollars' worth of beautiful equipment, but when I ask, "Where are your standards?" I am shown a ten-year old set of Johanssen gage blocks, or a standard cell that was last sent to N.B.S. in 1955, or I am told, "We crosscheck all our work in our own laboratory."

At Aerojet we recently carried out a designed experiment, using the statistical techniques of Analysis of Variance, to compare the results of pull tests on samples of stainless steel. We sent samples to two commercial laboratories. Without going into details, let me say that the results of tests made by these two laboratories were statistically different at the 1% level of significance. This is not to say that either one of them was wrong; but the question arose, "Who was right? What was the true tensile strength of the sample specimens?"

If either of these laboratories had shown evidence of good calibration on their equipment, we would obviously have accepted its results as more likely to be correct.

Since the facilities in the Los Angeles area for assuring accuracy and compatibility of measurements through sound chains of calibration are so inadequate, most defense contractors must send their equipment to the National Bureau of Standards. Besides overloading the Bureau, this involves from three months to a year of waiting while equipment is at the Bureau, plus serious risk of damage in transportation.

The American Society for Quality Control could make another major contribution to the defense effort by encouraging the establishment of adequate commercial calibration facilities in various geographical areas throughout the United States.

Now, how about the smaller companies, the sub-subcontractors and vendors? Aerojet has 2500 of these, most of them small. How do these people, who can not afford expensive laboratories, assure us of the accuracy and compatibility of their measurements? There is simply no way they can do this. And yet, Aerojet is contractually responsible for the quality of the work they do for us. Here we are, a major producer of missiles, with no way in which our vendors can honorably fulfill their commitments. The only thing we can do is what Napoleon is said to have done at one time when he was crossing the Alps with his army. He came to a deep crevasse, which appeared uncrossable. He passed the word on down through his generals and captains down to the corporals, "You have half an hour to get across." One very intelligent corporal found the answer quickly. He lined up his men on the edge of the chasm and ordered,

"Tenshun!! Fall out and form on the other side in fifteen minutes."

Most companies, especially the smaller ones, are in the position of the corporal who had to give an impossible order. These people cannot honestly meet the calibration requirements for government contractors.

The purpose of my talk today is to call these needs to the attention of ASQC and to suggest - in fact, to urge - that the American Society for Quality Control take an active part in developing adequate facilities for making measurements that tell the truth, because truth-telling is the heart and soul of Quality Control and Reliability. All of us together can accomplish several worthwhile objectives:

First, Support the National Bureau of Standards, especially in its budget requests to Congress and in public education on its vital role and functions as to National Defense.

Second, Encourage the establishment of adequate commercial calibration laboratories, in or near major metropolitan areas, by surveying needs, publicizing the profits accruing from assured measurement accuracy in Industry and emphasizing the calibration requirements now being placed upon contractors and subcontractors by the Armed Services.

Third, Grasp the opportunity, as an organization, to broaden our interests and our membership by taking an active part in this phase of Reliability and Quality Control which promises to grow like Jack's beanstalk during the next few years.

VIDEOSONIC INSTRUCTIONS RAISE QUALITY STANDARDS

David A. Hill, Manager
Semiconductor Division, Hughes Aircraft Company
and
John J. Tamsen, Quality Manager
Communications Division, Engineering Laboratory
Hughes Aircraft Company

HUMAN ERROR AS ONE CAUSE OF FAILURES

Analysis of the basic causes of failure in electronic equipment points to human error as a problem. The nicked wire, the solder blob, the forgotten process control, the incomplete inspection, the bungled maintenance routine--we are all familiar with them. One of two things has happened. The human being at fault has not known what he needed to know or he has been careless or forgetful. Our stubborn tendency to do the wrong thing is summed up in the bitter phrase, "When all else fails, follow directions".

There are many opportunities for human weaknesses to affect otherwise satisfactory equipment. These range from the inadequate description of requirements to misuse of equipment after it is in operation. The techniques to be discussed appear particularly useful for, but are not limited to, the prevention of human error during:

1. Manufacturing assembly operations
2. Inspection and test
3. Operations indoctrination and training
4. Field checkout and maintenance
5. Troubleshooting of malfunctions

FACTORS THAT CAUSE ERRORS

There are several factors which cause errors to result from human activity. Some factors for assembly operators and inspectors are:

1. Remembering long lists of part numbers
2. Identifying color coded parts accurately
3. Tracking through complex wiring instructions and diagrams
4. Carelessness with wire clippings, solder splashes, nuts, bolts, etc.
5. Incomplete or inaccurate assembly and soldering practices
6. Design change mental implementation
7. Distracted attention while working

Other factors are important for test and troubleshooting. They are, in part:

1. Long set-up and calibration routines
2. Safety precautions (doing things in the right order to prevent damage)
3. Numerous values and tolerances to check
4. Remembering special techniques needed to diagnose troubles
5. Losing track of specification changes

LCS 352:70:000

ATTEMPTS TO PREVENT OR DIMINISH HUMAN ERROR

Automation

The problem of human error can be tackled by taking the human out of the operation, as is attempted in automatic processing and testing. It will be some time before an automatic factory can handle the more complex assembly operations.

In electronics, the human being is likely to be a part of the production picture for quite some time. True, there are reports of methods for eliminating him--auto-assembly, deposited layers of film, monolithic circuit structures, etc. One needs only a moderately good memory to realize that hopeful announcements have been made at intervals for at least the last fifteen years. Like push-button warfare, this revolution is coming, but slowly. As for the operational and maintenance phase of an equipment's life, we must first accomplish fool-proof and perfectly reliable gear before human error is eliminated. A technique tailored to man's needs and weaknesses will have continuing use.

Improved Communications

Suppose we were to devise a system that would accomplish the following

1. Provide complete instructions developed by experts and checked out by the most experienced workers,
2. Communicate these in the most easily understood forms,
3. Release the workers' hands so that their sole function is to perform work tasks,
4. And most important-- give each instruction at the exact moment it is needed.

Is it not reasonable to expect that a dramatic reduction in human error would result if there was such an improvement in communications?

A lack of adequate communication is one of the greatest barriers to accuracy, reliability and efficiency. It is most difficult for the highly skilled engineer to convey to the operator all the necessary information in simple, easily understood language via the medium of written planning. (See Figure 1)

Ideally, we would have the engineer, the methods and quality experts, and the foreman sit with the individual production worker to guide him toward perfection. In recognition of this communications problem, Hughes established a special department to study and implement advanced manufacturing techniques. In the course of these studies, an entirely new concept in manufacturing systems was instituted. We call this new concept Videosonics.

Videosonics is the controlled visual flow of basic knowledge and experience through men, material and equipment. The Videosonic system is one of the most powerful tools ever developed for communicating intelligence in such a manner as to assure accuracy, guarantee standardization, and facilitate the understanding of the most complex data to individuals from all walks of life and of varying abilities. (See Figure 2)

Videosonic Equipment, Development

In recent years, we had used tape recorders alone as a means of communicating to our test technicians the sequenced test procedures and trouble-shooting hints. This not only cut down on the time required for a test, but allowed for better utilization of manpower; test technicians could then be checked

FIGURE 1 -- Written planning, blueprints, parts lists, assembly instructions, etc., contribute to confusion and errors.

FIGURE 2 -- Production line workers using new Videosonic systems no longer require written planning and the numerous written documents formerly required.

out on equipment in a fraction of the time necessary with printed procedures. Most important of all, hints on troubleshooting helped product reliability by reducing the number of trial and error parts-substitution, soldering and unsoldering of components, and number of on-cycles of the equipment under test. (See Figure 3a)

While tape recording techniques were being developed for testing, the use of 35 mm colored slides showing parts location, color coding, examples of good quality in workmanship, etc., were being developed for use with printed step-by-step assembly instructions. (See Figure 3b)

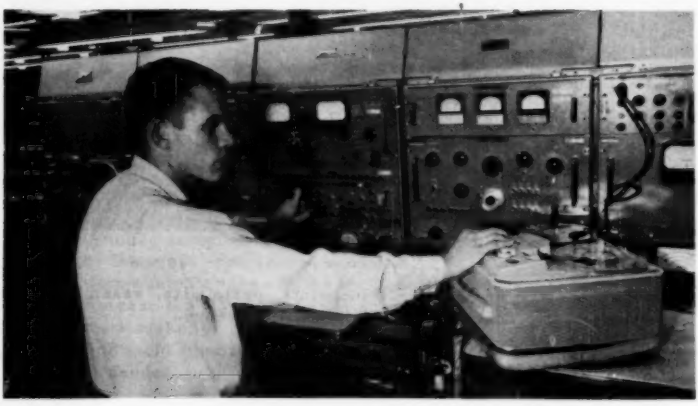


FIGURE 3a -- Early in the development of the Videosonic method, test procedures and troubleshooting data were recorded to improve utilization, aid the test technician in locating problems, and improve product reliability.

FIGURE 3b -- An assembly work station with slide projection to indicate a view of the product which the assembler ought to see on the part she is assembling.

The marriage of the two, tapes and slides, for a single process in test, inspection and assembly was inevitable. The tapes communicated to the assembler, the inspector, and the test technician the sequenced procedure/instruction while the slides showed what the individual ought to see on scopes, meters, product, test positions, etc. However, while the commercially available slide projectors and tape recorders have been developed to a point of high efficiency, they are not wholly adequate for our needs. This led to the development at Hughes of special Videosonic viewer-tape playback equipment designed for specific applications. We did continue to use the commercially available magazines for 35 mm slides and tape cartridges. The unit shown in Figure 4 is one of the many types of audio-viewers developed by the several divisions of Hughes.

FIGURE 4 -- A late model audio-viewer showing simple operating front panel control.

The Videosonic System

We next experimented to see just what the system would do for us. First, Videosonics required improved planning and programming which enabled us to break down into a series of discrete steps a process to be followed in a definite order. Normally, as more complex equipment was produced, more highly trained test and assembly personnel were required. The over-all Videosonics system enabled our company to use less highly trained personnel to a greater extent than would normally have been the case.

Second, pre-planning of fabrication functions was improved by the departments involved. Before the first article was produced, Quality Engineering, Test Engineering, Production Engineering, Methods and Time Study groups, line supervision, Planning and Videosonic section, and quite important, the most experienced line workers participated in preparing audio-visual slide and tape packages to cover the step-by-step operations. (See Figure 5) Note that this is done before the first article is made. The experienced workers cooperate fully to develop fabrication techniques, using shortcuts and assembly skills normally known only to the worker.

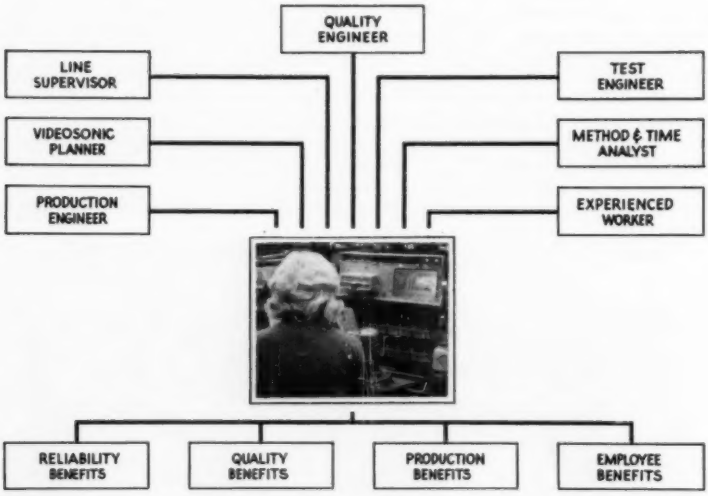


FIGURE 5 -- Components of Videosonic System

Third, there is an improvement in procedures and operations because Videosonics captures, accumulates and retains technical know-how. Because of the system, technical know-how is arranged in appropriate combinations and is finally communicated specifically and sequentially at the proper time to each person performing a task. When once prepared, a presentation may be used on long or short run projects, may be set aside for long periods of time, and when needed again may be used without the long period of indoctrination normally required to retrain a worker.

We have noted many improvements due to the implementation of Videosonics. They fall within four general areas:

1. Reliability improvements
2. Quality improvements
3. Production improvements
4. Improvements in employee relations

Detailed data and examples of the above improvements have been reserved for the section of this report entitled, "Benefits and Gains".

Specific Importance of Quality Engineering in Videosonics

We have mentioned earlier that the worker uses the projected view from the slide as a method of comparison. It is important, therefore, that each slide be accurate not only of component location, but also as to workmanship detail. Certainly, a poorly dressed wire or a sloppily oriented part in conflict with the oral instructions could lead not only to confusion in the worker's mind, but could result in a continuous flow of defective parts. Therefore, before the slides are used or the tapes prepared, quality engineers review the scripts and scenes to assure that they meet quality standards. A special Quality Engineering approval stamp is placed upon each slide so that it is visible when viewed on the screen. Only approved slides may be used. This

assures the worker that what she looks at is an approved view of the method and workmanship to be expected of her. The stamp appears under the cover of the Fidelipac cartridges for the tape recordings. Step by step visual inspections are built into the tape and slides as part of the production instructions.

APPLICATIONS OF VIDEOSONICS

Each of the fourteen operating activities within Hughes Aircraft Company has in some manner implemented the Videosonics System. The Falcon Missile at Tucson, Arizona, is assembled, inspected, tested, and supported in the field through Videosonics. The Fire Control Systems manufactured at El Segundo are likewise supported and in addition the depot, operated by our Field Service and Support Division, uses Videosonics in repair, modification, and checkout. Videosonic experiments in our own Communications Division have been very successful. The various activities concerned with commercial products have implemented the system in these and other functions, such as process control, safety in handling chemicals, training, etc. and over-all, the Research and Development Laboratories at Culver City have designed new equipment with Videosonics installed to assist field operators in handling the more complex field checkout equipment.

There is no apparent end to the system's utilization. Our training departments, sales activities, employment departments, process control and fabrication operations, libraries, central document control areas, purchasing departments, and many others are currently using or are planning for future use of the system. One unique innovation, starting during the summer of 1959, was the taped instructions and slide views of forms for filling out employment questionnaires and security clearance applications; thus replacing a somewhat complicated instruction sheet.

In the area of customer use, the system boosts tactical readiness by supplementing the present technical capability levels of military support personnel. The system lessens the effect of personnel turnover. It assists in isolating and identifying authentic design or manufacturing deficiencies for effective corrective action, thereby accelerating squadron activation. Videosonic devices reduce military operating and maintenance costs.

Several months ago, a technical order was issued for the Falcon Missile. The complexity both in number and type of changes caused no end of confusion. A special team was sent to the depot to assist. This helped for a while. On a chance that Videosonics might be the answer, a special package was prepared on this difficult T. O. and sent to the depot with an audio-viewer. The special team from Hughes tried it out successfully and when ready to let the Air Force experienced technician attempt the task, learned that he was in the hospital. Another technician was selected, one who had no knowledge of the Technical Order, and on his first attempt using Videosonics, accomplished the task in far less time than the experienced technician did without the benefit of the special package. This emergency answer to a need has resulted in the complete utilization of the equipment for the balance of the Technical Order and for others to follow.

BENEFITS AND GAINS WHICH WE HAVE EXPERIENCED

The benefits and gains realized since the implementation of the Videosonics system have fallen into four basic groups. They are:

1. Reliability benefits
2. Quality benefits
3. Production benefits

4. Employee relations benefits

Many of the benefits will overlap between groups. However, for convenience in discussion, we will observe them as associated with but one group. Let us look at these benefits in the above order.

Reliability Benefits

We have discovered through continuous sampling and statistical evaluation that reliability, especially in the Falcon Missile, is dependent upon a unit's first acceptance through test. Components which fail when first tested generally require rework. During rework, parts must be unsoldered, removed, and new parts resoldered. Such resoldering weakens the printed circuits and adjoining connections to such an extent that the probability of a subsequent failure is increased. Videosonics has provided a method for improving workmanship. As a result fewer defects are noted during test which are traceable to workmanship faults. Before Videosonics was applied, between fifty to sixty percent of all electronic packages submitted passed their first test successfully. After Videosonics, between seventy-five and eighty-five percent passed their first test. On the most complex chassis in the Falcon Missile, rarely did better than fifty percent pass the first test. After Videosonics, the rate had increased to an average of eighty-four percent. An exhaustive evaluation was made of this improvement. All squawks were analyzed and assigned as either a component failure or a workmanship defect. Component failures were defined as parts which failed because of hidden defects or inherent faults which were destined to show up regardless of the care exercised by the worker. Workmanship defects include all faults in wiring, soldering, installation, and any other problem built into the units during assembly at Hughes. The final analysis before, after and during the implementation of Videosonics, is shown graphically in Figure 6.

Component failures have remained relatively constant throughout the control period. Workmanship defects have shown a rapid decrease during the period of implementation of Videosonics, and have leveled out to a point which is well within any calculable control limits. As a result, we have generalized a theory that if a unit will not pass its first test the fault lies with a component part. Each failure is carefully analyzed in order to increase the reliability by procuring better parts or helping our suppliers build more reliable components.

Other divisions of the company are experiencing similar results. For example, on our Fire Control Systems the average test defect per unit rate has decreased from 1.64 to 0.28, after Videosonics was applied. The decrease in defect rates is further reflected in an average savings in test time of about one hour thirty minutes per unit. On one unit, formerly having the highest defect rate, the average test defects per unit fell from 5.02 to 1.50.

Quality Benefits

Our quality benefits are measured in terms of reductions in defects per unit or per standard hour as recorded by inspection. Videosonics has established a standard approach in assembly. The assembler sees the location of parts, the routing of wires, etc., in color exactly as they should appear on the assembled part. They hear and see that the black band of a capacitor is on the right. They hear and see that the red wire goes to point X and the yellow wire goes to point Y. As indicated in the section above on reliability, reversed wires and incorrectly oriented parts seldom occur. With Videosonics at inspection, the inspector is told to look for certain parts and wires and their relation to others. They also see, for comparison, the colored

slide and the part which is being inspected. This combination of activity, Videosonics at assembly and inspection, has resulted in an average defect per unit rate decrease from 1.2 to .115, less than ten percent of the former rate. On one unit with 347 solder connections, about 150 parts, and with a 21 wire harness the average defects before Videosonics was 0.6 per unit. After Videosonics, and during the entire period to date, there has not been one single workmanship squawk recorded.

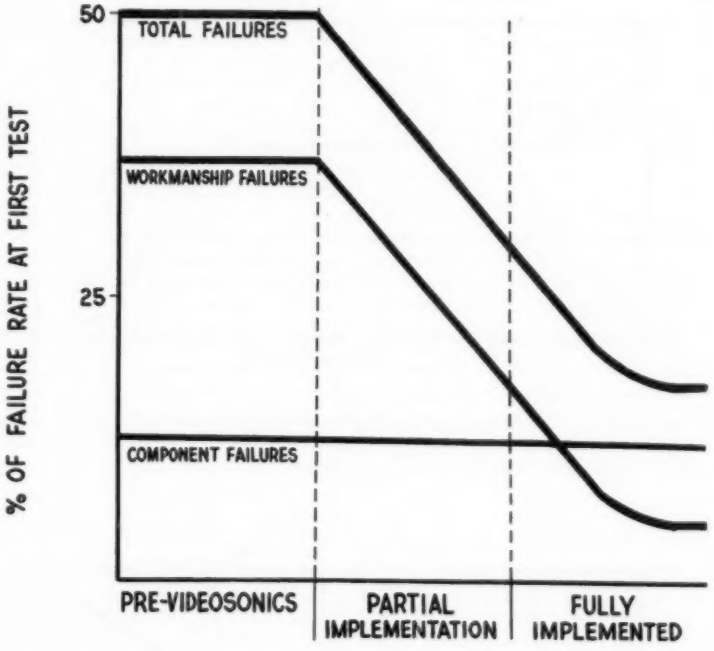


FIGURE 6 -- Chart showing relation of causes of failure discovered during first test of electronic components in Falcon Missile.

Figure 7 shows the graphical relation of defects before, after, and during the initial implementation of Videosonics. The sharp rise noted during the period when inspection first used Videosonics is indicative of the more thorough inspection which discovered defects normally found during test, or not discovered at all. When once located, concentrated efforts were made to eliminate recurrent defects by planning corrective action in assembly Videosonic presentations.

Production Benefits

Production benefits are best shown in the improved utilization of the available manpower. We were operating before Videosonics at a rate of 47%. The average utilization today on twenty-five different manufactured items is 90%. A typical assembly line at our Tucson facility, formerly had 7 assemblers, one rotating substitute assembler for absences, one rework assembler,

one checker to review the items before inspection, and one and one-half inspectors. After Videosonics, the checker was deleted, there was no need for the substitute assembler, and because of the vast decrease in defects, the rework assembler could not be utilized full time, and the inspector dropped to one-half. Also, of significance, the number of assemblers required to produce the same number of items fell from seven to five. Figure 8 shows the steady climb in utilization rates for the several units.

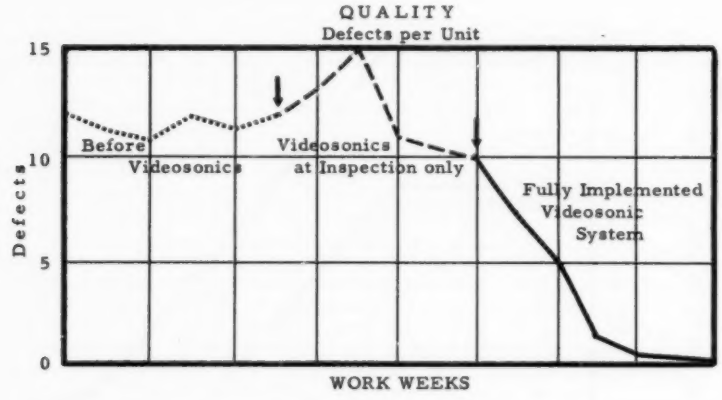


FIGURE 7 -- This chart shows the trends recorded in the reduction of defects before, during implementation, and after Videosonic assembly and inspection, as recorded at inspection.

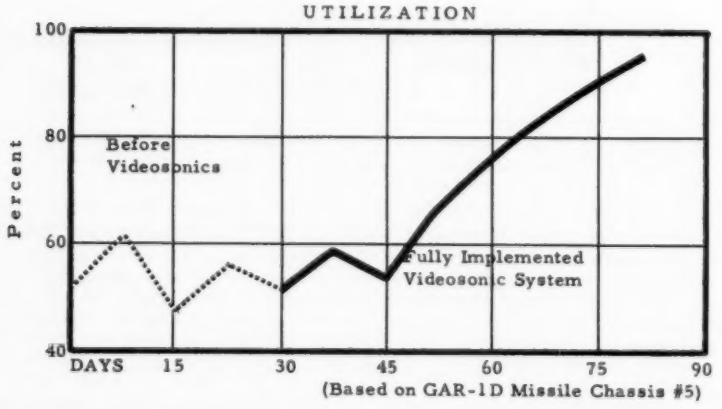


FIGURE 8 -- This chart shows the steady rise in the utilization rate experienced due to the implementation of Videosonics.

Employee Relations

There is an increase in the general morale with a more desirable attitude expressed by the individual employee. Direct measurement within this area is difficult to obtain, if not impossible. However, in general we have noted a decline in the absence rate of female employees, a marked increase in the number of new hires who stay over a longer period (a reflection of a decreasing turnover rate), etc. Of particular importance is the quality measurements of individual worker performance. Formerly, an assembly line as a whole was measured. After Videosonics was applied, each assembler was individually measured for quality performance; an advantageous feature of Videosonics makes this possible. Now workers with like abilities and speed can be relocated to advantage, that is, take slow workers from a fast moving line and replace them with faster workers so that the line may move at a constant speed. Through an extensive quality indoctrination program, the line assemblers have accepted the individual measurement plan and strive to improve their own records. We do not use this plan as a lever through which the worker is forced to do better or else. Rather it is a motivation to build pride in workmanship, a pride which the worker feels is important in his work.

The learning curve in training or retraining is a function of the time taken to become proficient enough to do productive labor at a point equal to or slightly less than that of the average production worker. The learning curve under Videosonics has gone up rapidly. New employees are reaching a point through the use of Videosonics in less than half the time formerly required without Videosonics, and within the same time are nearly twice as proficient. Employees are relocated on new assignments with but one day's practice on Videosonics and on the second day are as proficient as the regular workers. Normally, without Videosonics, a relocated worker would require about two weeks for the same results. Figure 9 shows various relations in learning time.

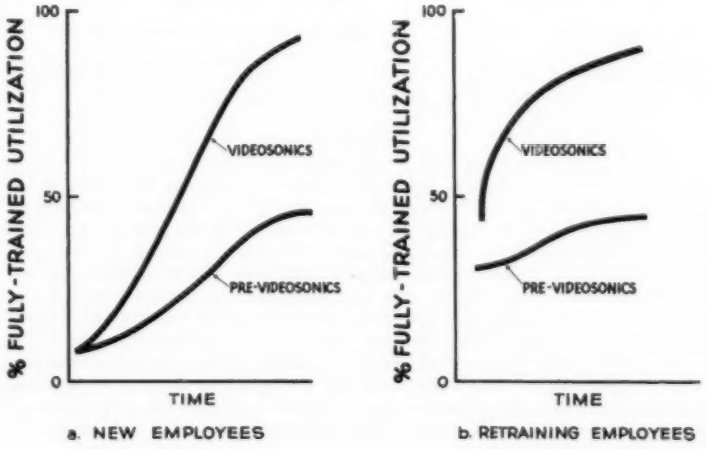


FIGURE 9 -- Relative improvements in learning curves for a. New employees and b. Relocated employees new to the function.

COST SAVINGS OF VIDEOSONICS

Extensive reports are being made upon the results of production using Videosonics. One study, comprehensive in nature, deals with costs expended and costs saved. Cost records show manhours saved and expended before and after Videosonics, equipment costs, material costs, plant costs, etc. In fact, every cost within the factory and administrative applications is recorded. The crossover point came in the third year. During the following years the savings exceeded the cumulative cost of Videosonic systems. Figure 10 shows first the crossover point and the relative points of cost savings and expenditures, and the components of costs at the point of crossover.

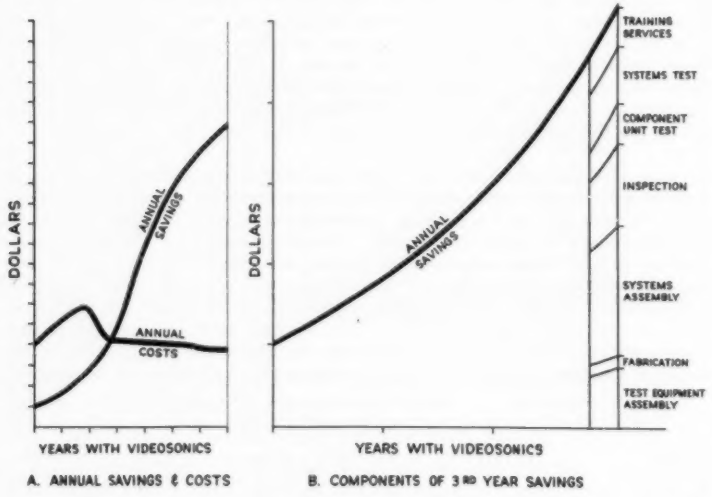


FIGURE 10 -- a. Crossover between costs saved and costs expended.
b. Components of costs saved at point of crossover.

SUMMARY

In this paper we have told of our experiences in Videosonics. We saw how human error in assembly, inspection and test is one of the greatest causes of failures in electronic equipment. Human error can, to a certain degree, be eliminated by automation. However, certain functions require the exercise of the human mind and hands. To these functions a better means of communication was the answer. One method devised at Hughes was the audio-visual means of communication, which when fully implemented became what we call Videosonics. Videosonics is a means of controlling the visual flow of basic knowledge and experience through men, material and equipment. In order to fulfill the need of this philosophy, we had to develop new manufacturing methods and special equipment. The equipment was designed so that the worker had merely to turn it on and proceed to work without having to remove her hands from the tools or item being made. The equipment was fully automatic.

We have applied Videosonics at assembly, inspection and test. We also have used the equipment in other non-productive areas, such as employment, training, record keeping, sales, retention of process techniques, and so on.

Figure 5 indicated the components of input and output of the system. The outputs of the system are in reality the areas where benefits have accrued. We noted them as being reliability, quality, production and employee relations. The costs incurred in the implementation of the system have exceeded the cost savings during the early phases but a crossover point was seen in the third year of operation, and from then on the savings have exceeded the cumulative expenses each year.

Videosonics has improved communications, reduced our costs, improved the ability of our company to move quickly to new market situations, and has greatly improved the reliability and quality of our products.

AUDIT TESTING

C. W. Carter, Industrial Management Consultant
Longmeadow, Massachusetts

Introduction:

The title may suggest to the reader a double interpretation. First, this article is a discussion of a particular type of testing. Secondly, it is a suggestion that all testing should be audited. Most managements will agree that it is worth a nickel to know that a dollar's worth of test results are correct.

What is meant by "audit testing"? According to Webster, an audit is a "general and formal examination". Used in this connection, then, it means a periodic examination of routine testing to assure that it is being done when and as specified. Duplicate samples, selected at random, are run by the routine Tester and the Test Auditor. If the results differ by a larger amount than can be statistically attributed to pure chance, a careful on-the-spot study is run to determine the reason or reasons for the difference. Such a program, in operation at the International Falls mill of the Minnesota and Ontario Paper Company, will be described here. Although slanted particularly at the paper industry, it is felt that these techniques have considerable applicability to other industries.

Background:

Such a program did not happen just by chance. Perhaps the roots from which this activity sprang lay in the question, "What is the function of testing?"

This has a threefold answer:

(1) To provide the operator with information which forms the basis for taking action to minimize or even forestall the manufacture of sub-standard product.

(2) To give both operator and technical personnel information for diagnosis and analysis. Although after-the-fact, this can be often useful in safeguarding against future troubles.

(3) Acceptance of material for further processing or shipment.

It is pertinent here to emphasize that the action taken and the analysis of trouble can be no better than the data upon which it is based. It, therefore, becomes extremely important that the data obtained be of sufficient accuracy to lead to the correct decision. Of equal importance is the requirement that the total amount of testing be within the budgetary limits.

There are, however, two circumstances - not necessarily peculiar to but found in the paper industry - which make sampling and testing a real problem. First, there is the difficulty of obtaining a random sample. For example, paper coming from the machine is wound until the roll reaches a certain diameter. It is then broken off and a new roll started. A random sampling plan might call for testing the 1523rd yard. However, this might well come somewhere between the beginning and end of a roll. The result

of this impasse is a practical compromise whereby the end of the roll is assumed to be representative.

Also, the sampling and testing of stock and furnish at the different stages of manufacture is traditionally set so that an inspector may have a fixed schedule. This routine type of control, although easy to administer, has an element of inflexibility and does not always readily adapt itself to changing quality needs.

The second peculiarity is that testing in a paper mill is carried on by two and sometimes three responsibilities. Many tests are performed by production personnel, although the majority are done by inspectors and laboratory testers. This different responsibility always gives rise to differences in sampling and testing procedures; indeed, sometimes in test equipment. It also tends toward slower communications, which hamper the effectiveness of a testing program.

These, then, offer challenges to a paper testing program, which must be met on some basis consistent with quality and cost if the testing effort is to be justified.

The Approach:

As a first step in solving any problem, it is advisable to get the facts. This was done by conducting a testing survey, which embraced all testing in the mill. The procedure here employed was to compile a list of tests by departments, using the form shown in Figure 1. This basic tabulation is composed of four general headings, each with subject matter as follows:

(1) Test

For each department was developed a complete list of quality characteristics tested. This was accomplished through talks with production supervisors, testers, inspectors and technical personnel. The survey included all production departments from the Wood Room through Shipping.

INTERNATIONAL FALLS PLANT
QUALITY CONTROL DEPARTMENT

Survey of Testing

Department _____

	Sampling				Testing			Results		
Test	Where	Frequency	By Whom # People	Time Required	Where	By Whom # People	Time Required	Who Uses Results	Action Taken	Effective For Control? If not - Why?

Figure 1 - Survey Testing Form

(2) Sampling

This heading describes the method of obtaining the sample. Under the sub-column headed "where" was entered the location of the sampling point. Next, the "Frequency" of sampling a process (i.e., once per shift, every hour, once a reel, etc.) was entered. "By Whom" and the "Number of People" helped to identify the responsibility for obtaining the sample. Finally, the "Time Required", although only an approximation, was useful in comparing the total time spent in sampling with the relative importance of a given characteristic. It should be here emphasized that no effort was made to "time-study" this activity.

(3) Testing

Here again, the location ("Where"), responsibility ("By Whom"), and the amount of effort ("Time Required") were entered. These allowed comparisons with the sampling activity, and just listing these side by side brought out several suggestions for strengthening, consolidating, and improving the effectiveness in specific cases.

(4) Results

Whereas the first three headings were susceptible of very factual treatment, this was more a compilation of subjective judgements. These were arrived at through detailed talks with production and technical personnel and by observing the use to which test results were put. Several parties might use a particular test and the "Action Taken" would differ according to the user. As a guide to the evaluation column "Effective for Control", a series of ground rules were established and discussed with the persons concerned. This classification as shown implied the alternative of "Not Effective for Control". During the course of the survey, this was moderated to the classifications suggested by the first two functions of testing (see section on The Background). These might be paraphrased "on-the-spot action" and "after-the-fact analysis". It is apparent that a test falling into the first classification will certainly fit the second. A test used for on-the-spot action (e.g., headbox pH) may also be used for analysis purposes later. However, the converse is not necessarily true. Beater Value, for example, is not used for making any on-the-spot adjustments, but is studied by operating personnel as being indicative of the general run of wood and quality of pulp. It was also stressed that classification under the second description did not imply the recommendation to eliminate or reduce the frequency of a given test.

During the course of this fact-finding phase, many changes were suggested by participating personnel. Some were implemented immediately, others were discussed in general supervisory meetings. Through careful study of the data, further opportunities for strengthening the testing were uncovered. Being somewhat less obvious, these required considerable discussion before arriving at satisfactory recommendations.

Results:

The testing survey proved to be a many-sided document. The broad recommendations were as follows:

(1) That certain sampling and/or testing procedures be changed to obtain more accurate results.

(2) That statistical process control limits be established in order that the tester may notify the appropriate authority when tests are beyond limits.

(3) That the frequency of certain tests be reduced. This can be accomplished in two ways: first, more efficient use of the test data (i.e., more information per test by using statistical techniques), and, second, adjusting the time between tests in accordance with the innate stability of the process.

(4) That certain tests be eliminated.

(5) That a program of audit testing be instituted to make the present control testing more effective.

Certain by-products accrued from the testing survey: A Manual of Testing was developed, a typical page from which is shown in Figure 2. Composed by departments, this has proven a valuable reference for many. Responsibilities are herein formalized, which makes it useful to supervision. The lines of communication are strengthened and this tends to hasten action taken on out-of-control conditions.

It was found that definition of test procedures was helpful in lessening the variation due to sampling and testing. In one case, the need for new testing equipment was demonstrated and the apparatus obtained. Changing the location of sampling points, as well as testing points, helped to make certain tests more effective for control. Management has better assurance that the basic information for judging conformance to quality standards is reliable.

The final recommendation led to establishing the formal position of an audit tester in the Control Laboratory. The audit tester is an experienced laboratory man with a sound knowledge of all the standard sampling and testing procedures. His only additional equipment is a table of statistical "duplication" limits. It is his function to make the rounds of the mill, covering a major share of the testing being done, and to run comparison samples with the tester. Whenever practical, the operators take duplicate samples throughout their shift. These are selected at random by the audit tester, who makes the test.

If the results are no farther apart than the "duplication" limits, it is considered an agreement. If the two results differ by more than the statistical "duplication" limits, the audit tester immediately reviews the test procedures followed with the standard test procedure. If the reason for the difference is not readily apparent the audit tester again obtains and runs another sample with the routine tester.

The duplication limits may be calculated in a number of ways depending on the sample and testing procedures followed. Two examples of how duplication limits may be determined are shown in Tables I and II.

Sulfite Mill

No.	Test	Sampling Where	Sampling Frequency	Testing Where	How Results Are Used
ACID MANUFACTURING PROCESS					
1.	% SO_2 Gas Test TAPPI 603 M - 45	From gas line off coolers	Once every two hours	Acid Plant	This test is used for checking the automatic Leeds & Northrup SO_2 recorders.
2.	Temperature of gas, strong tower and weak tower M & O Method		Once per hour	Acid Plant	The Acid Maker controls the temperature of the gas, the strong tower, and the weak tower to regulate the speed of the gas absorption in the acid making process.
3.	Acid test, total, free, and combined SO_2 TAPPI 640 M - 45	From strong and weak towers	Once per hour	Acid Plant	This test tells the Acid Maker how much lime rock to use. He adjusts the amount of water and gas going to the towers with this information.
4.	Baume on acid in strong tower and acid in weak tower. M & O Method	From strong tower and weak tower	Once per hour	Acid Plant	The Acid Maker controls the strength of his acid in the towers with these test results. It tells him the concentration of his acid.
DIGESTER COOKING PROCESS					
5.	Chip Moisture M & O Method	Sample taken in Chip Bin	Once per cook	Chip Room	The cook adjusts the volume of the chips he puts in the digesters with this information.

Figure 2

TABLE I - Duplication Limits Calculated from Average Range Between Test Results Obtained on Two Samples Selected from the Process at the Same Time.

Time	Sample 1	Sample 2	Range
10:30	3.92	3.87	0.05
1:05	4.03	4.00	0.03
4:20	3.92	3.92	0.00
6:30	3.92	3.95	0.03
9:05	3.97	3.92	0.05
12:20	4.00	4.00	0.00
2:30	3.90	4.00	0.10
5:05	4.00	3.92	0.08
8:20	3.97	3.97	0.00
10:30	4.03	4.00	0.03
		Total	0.37

Average Range ($\bar{R}$) = 0.037

The maximum range permitted between an Audit and Routine Testers results on samples taken together is set at:

$$D_4\bar{R} = 3.3 \times 0.037 = 0.12$$

(99.7% probability level)

TABLE II - Duplication Limits Calculated from Variation in Test Results Obtained on a Number of Duplicate Samples by One Tester.

	X Test Result	X^2 Test Result^2
	15.2	231.04
	15.0	225.00
	14.7	216.09
	15.0	225.00
	14.6	213.16
	14.6	213.16
	14.4	207.36
	14.7	216.09
	14.4	207.36
	14.7	216.09
Total	147.3	2170.35
Average	14.7	

Test Standard Deviation (σ) = $\sqrt{\frac{2170.35 - \frac{(147.3)^2}{10}}{9}}$

= 0.263

The Standard Deviation for the Difference between two testers is:

$$\sigma_d = \sqrt{(0.263)^2 + (0.263)^2} = 0.37$$

The 99.7% probability limits for agreement are:

$$\pm 3\sigma_d = \pm 1.1$$

It is important that the audit tester not be restricted to a definite schedule. There are two reasons: First, the routine tester should have no way of anticipating the arrival of the audit tester and, second, there should be sufficient time for the audit tester to review the sampling and testing procedure should a difference arise. The audit tester work sheet is shown in Figure 3. This forms the basis for a monthly summary to management.

It should be noted, in connection with the establishment of the audit testing program, that it was done with no increase in the testing manpower. This time was made available by rescheduling the testing assignments among present personnel.

Summary:

From the basic question, "What is the function of testing?", came the fact-finding phase of a testing survey. Perhaps its most important outcome was the initiation of an audit testing program in the mill.

For a very nominal investment in time and through a systematic approach, a sound program of audit testing may be developed. Since errors in process control often have serious consequences in off-specification product, it is essential to know that all tests are reliable. Further, relatively little is known of the interrelationships of characteristics measured early in the process with the quality of finished paper. Only through meaningful tests may these relationships be accurately determined. It appears that audit testing gives assurance of such reliability at very low cost. Its value has been demonstrated - try it.

AUDIT TEST FORM

Test Audited _____ Department _____

Audit Tester _____ Period _____

Control Limits _____

Test Number													
Date													
Tester*													
Routine Test (X)													
Audit Test (Y)													
Diff. (X-Y)													

† = Recheck

Difference

* Routine Testers Code: A = B = C =
D = E =

Remarks:

Figure 3

VARIATIONS OF MECHANICAL PROPERTIES IN ALUMINUM PRODUCTS

W. P. Goepfert, Head, Statistical Analysis Department
Aluminum Company of America, Pittsburgh, Pa.

INTRODUCTION

Mechanical properties of aluminum products is a subject that receives much attention both by the producers and consumers of these products. Producers are constantly striving to develop new alloys and tempers that possess optimum combinations of properties to satisfy concepts of applications conceived by design engineers. The objective is not merely one of producing a product with high tensile properties. Other characteristics such as formability, resistance to various types of corrosive environment, fatigue strength, compressive strength, shear strength, freedom from residual stresses that result in warping when the product is machined or selectively reduced in size by chemical methods, etc., are involved in developing an optimum combination of properties. Variability in properties with respect to location of test specimen is another important fact that invites attention. In thick plate, for example, how much bias in properties can be expected across the thickness? How do the properties on specimens taken from the middle plane of the plate compare with those from a plane midway between the center and surface? How do properties of surface specimens compare with those from the center or midway plane?

Direction of testing is another variable that must be evaluated. Specimens cut with their axis parallel to the direction of final rolling or extruding will usually possess different properties than specimens cut from the same relative location in a given piece but with their axes parallel to the width of the product. Another level of properties might be obtained on specimens cut with their axes parallel to the thickness dimension of the product.

Properties at temperatures or after subjection to temperatures other than ambient are also of great important and must be known for certain types of applications. Naturally, the properties of some alloys are affected more by high temperatures than others. Similarly, properties of some tempers of the same alloy are not affected to the same degree by elevated temperatures. In general, low temperatures do not adversely affect the properties of aluminum alloys. Development of some of this knowledge is a matter of research laboratory studies. Others are the result of statistical studies of data developed in routine testing or special test programs at the plants. In this paper, some of the programs being carried out with respect to plant data will be discussed.

SPECIFICATION REQUIREMENTS

First those properties which are normally specified as requirements of specifications defining particular aluminum products will be discussed, as well as the underlying philosophy concerning the significance of these guaranteed values, the methods used in arriving at the proper values, and the system that has been established to keep a constant check on our performance against established guarantees in order to see that current capability continues to be compatible with them.

The properties commonly specified as guarantees in aluminum alloy specifications are tensile ultimate and yield strengths, elongation and, for some products, Brinell hardness. The requirements are established as maximum and/or minimum values to be met

by standard test specimens cut from the product with their axes in a prescribed direction with respect to the product's length, widths or thickness. For some products, even more specific requirements are established for location of the test specimen in the product. The number of tests to be made per lot are specified usually on the basis of one for every so many pounds or feet of the product in the lot. This might mean one test per lot, or several tests per lot. In general, the acceptance requirement is on an attribute basis, i.e. the prescribed guarantee value must be met by the test on a go-no go basis rather than on a variables basis that would take into account the actual magnitude of the test values obtained and the variability of these values within a lot. From a statistical viewpoint, it can be seen that with such relatively small sample sizes considered as an attribute sampling plan, the probability protection level afforded by the prescribed test procedures is not very high.

In addition to the inherent weakness of the prescribed sampling procedures to give a high probability assurance of compliance of all the product in a given lot with guaranteed properties, retesting is often permitted on the basis of two tests for one in the event of failure of the original test. In this event both the retest values must meet the guarantees on an attribute basis in order for the lot to be considered acceptable. This further increases the probability of failure to reject lots containing some material outside of guaranteed limits.

This is not intended as a criticism of the sampling procedures of the specifications. The question is one of economics and philosophy. Economically, a sampling plan that would in itself provide high assurance of practically all the material in the individual lots conforming with guaranteed values would be prohibitive. Under normal production conditions, the physical test equipment and personnel are pressed to keep up with the amount of routine testing required under these present sampling requirements. It might be of interest to know that the average monthly load on testing facilities at one of our larger Works is of the order of 20,000 tensile test specimens. This is close to a quarter million tensile tests performed in one year at just one of our production facilities. This should be more than ample check on the process controls as reflected in mechanical properties.

In view of the above, the strength of assurance in mechanical property conformance must lie in the establishment of guarantees that represent the capability of the product and in the producer's system of control.

We feel that establishment of guaranteed mechanical properties on the basis of sound statistical analysis of data representative of the process capability is the only solid foundation on which to build a system of high conformance of material shipped. In general it is current practice to consider that the minimum guaranteed value should represent that value outside of which it can be predicted with 95% probability that not more than 1% of accumulated data from material produced by the given standard practice will fall. Because of the destructive and costly nature of mechanical property testing and inherent variability within lots and within pieces within lots, it is not practical to rely on testing to separate out material that is below guaranteed values. The major emphasis must be on the proper establishment of guarantees based on process capability and on processing controls to assure controlled production. Final property testing then assumes the role of a spot check for evidence of nonconformance rather than a high probability check for conformance. This distinction is somewhat analagous to the alpha and beta risk concept in statistical tests of significance.

WITHIN LOT VERSUS ACROSS LOTS VARIABILITY

There have been times when certain consumers have proposed that a premium strength classification of a given alloy and temper of a product should be offered by aluminum producers. These suggestions usually have involved the proposal that the premium strength lots could be obtained by selection based on test results obtained on representative samples from the lots. These proposals have not been readily accepted by the producers because of the fact that variability of mechanical properties within a lot of material is often of the same order of magnitude as that across an accumulation of a large number of lots. This is to say that the process is in control and that all the production lots of a given product can be considered to be from a given population of mechanical properties. An example is presented here to indicate this statistically. In order to develop a measure of the variability within lots of one sheet product, control charts were established for 0.064 inch thick Alclad 2024-T3 sheet for a period of time. The procedure was to select three random sheets from each lot, test a sample from each sheet, and plot the average and range of the results for each lot so tested. This procedure was carried out over an extended period of production and indicated a good state of control. For comparison, the standard deviation of all production data for a six month period on the same alloy and temper for all sheet thicknesses was reviewed. The following table shows this comparison.

TABLE I
Within Lot Vs. Across Lot Variability
Alclad 2024-T3 Sheet

	Within Lot		Across Lot
	$\bar{R}$	σ	σ
Tensile Strength, PSI	2100	1,240	1,300
Yield Strength, PSI	1900	1,122	1,500

The above type of comparison indicates the impracticability of realistically segregating premium strength lots by selection based on test results.

CENTRAL OFFICE SURVEILLANCE OF
PLANT MECHANICAL PROPERTY PERFORMANCE

In addition to the system of standard practices for production of a given product which are designed to produce a product that will consistently meet its quality requirements, including mechanical properties, and the routine surveillance of performance at the plant level, Alcoa has a system by which all routine mechanical property data from its producing facilities are summarized and submitted to the Statistical Analysis Department of the Metallurgical Division of Pittsburgh. These data are analyzed statistically to determine the capabilities of the several producing plants for a given product with respect to existing guarantees or as a basis for establishing firm guarantees on newer products as they develop.

Those practicing in the field of industrial statistics are certainly aware of the difficulties that can arise in analyses of accumulated data on this type due to biases unless considerable care is taken to see that the data to be analyzed are representative of the process. This requires the establishment of some very definite rules as to the composition of the masses of data submitted for analyses. Such rules have been established, based on our

experience, to assure within practical limitations that the data submitted from the plants can be considered representative of the product when producing under the established standard metallurgical practices. These rules and other details governing the routine mechanical property data reports submitted to the Central Office are established in our Standard Practice for Mechanical Properties Summary Reports.

The following excerps from this Standard Practice Manual will indicate the type of rules established to assure that the data submitted is representative of the process and not merely an unbalanced compilation of data on lots that met the established guarantees.

The report shall include the following data:

1. All original tests on material produced by standard practices and tested in full section or as standard type specimens taken from the standard test location for the particular product when

   (a) Material was produced in prescribed manner for the specified alloy and temper. Only positive proof of improper processing is cause for exclusion of original data.

   (b) Break is between but not through the scribed or punched guage marks.

   (c) Fracture of failing test specimen does not exhibit excessive discontinuities such as dross, blow holes, flakes, etc. to the degree that justifies replacement test.

   (d) Data are from random samples from the lots and not samples selected for testing because properties were considered questionable based on some prior non-destructive test such as hardness test.

   (e) Number of tests per lot is not more than usual frequency of testing for the respective products. When more than the usual number of original tests per lot are made for special purposes, select therefrom, in testing sequence, the usual number of tests and include these only.

   (f) Specimen is tested under proper procedures on properly functioning equipment.

A similar list outlining data not to be included in the report is also given. This excludes data such as that from lots that were reheat treated, reprocessed in any fashion or retested as a result of failure of original tests. Compliance with these rules results in reliability of data for the purpose of establishing process capability.

The Mechanical Property Summary Reports are submitted on a semi-annual basis on a staggered schedule so that all plants producing a given product will report their data on that product accumulated over the same calendar six months of production. These reports are in the form of frequency distributions for the data within standardized size groupings of the product. They show, in addition to the frequency distribution, the number of tests, the

average and standard deviation of the data and the guaranteed values applicable. Table II is a sample sheet showing the form in which these data are reported.

TABLE II

MECHANICAL PROPERTIES SUMMARY

2024-T4 PLATE 0.250— 0.500

TYPE OF TEST	OPERATING LIMITS MAXIMUM	OPERATING LIMITS MINIMUM	TEST RESULTS CELL	TEST RESULTS FREQUENCY	TEST RESULTS AVERAGE	TEST RESULTS STD DEV	TOTAL NO OF TESTS
TENSILE		64 0			67 109	1 162	470
			72	1			
			70	3			
			69	23			
			68	82			
			67	149			
			66	149			
			65	47			
			64	13			
			63	3			
YIELD					46 273	1 656	470
		40 0					
			51	1			
			50	9			
			49	18			
			48	41			
			47	90			
			46	109			
			45	102			
			44	65			
			43	28			
			42	5			
			40	2			
ELONG 0.250—0.500					19 402	1 214	470
		12 0					
			24	1			
			23	1			
			22	11			
			21	50			
			20	114			
			19	126			
			18	122			
			17	43			
			16	1			
			12	1			

On receipt of these reports in the Central Office, the statistics from them are plotted on standard charts showing the historical performance of all plants producing the pertinent product. Figure 1 shows a sample of such a chart. You will note that in addition to the identifying information, there are plotted the average and standard deviation for each plant. Symbols are used to identify the plants. The number of tests making up the pertinent statistics are shown for each time period and in practice the number of these that actually failed to conform with established guarantees are indicated in parentheses alongside the number of tests. In addition to recording failures to meet guarantees, a similar record is kept on the back of the chart of performance against established Aircraft Structural Design "B" values. These are values above which at least 90% of accumulated test data are expected to fall, and are used for design purposes by the aircraft industry under certain restricted situations. It might be of interest in visualizing the magnitude of this program to know that currently close to 15,000 of these charts recording mechanical property performance are maintained. Each chart covers a given property for a given alloy, temper, product and size grouping.

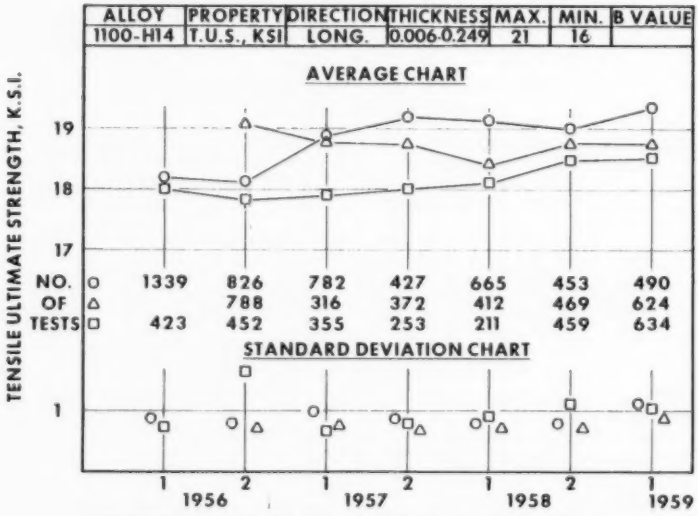


FIGURE 1 - MECHANICAL PROPERTY CHART - SHEET

What are these charts used for? They are reviewed for any evidence of excessive failure to meet established guarantees or "B" values. Similarly, they are reviewed for any evidence of greater capability than current guarantees or "B" values. In either event, such cases are brought to the attention of cognizant personnel and further investigation initiated to see whether changes in existing guarantees are indicated. These charts are carefully reviewed for any evidence of noteworthy differences in performance of individual plants which warrant investigation to determine causes. The standard deviation charts are watched for evidence of too much variability in data within a size group. This might be indicative of the desirability of further breakdown of size groupings for guar-

antee purposes that would permit of favorable changes in guarantees for a portion of the original size groupings. These are only some of the purposes served by the charts. In general, they are very useful in readily showing what is happening with respect to mechanical properties of various products and in pointing out areas for further investigation.

RELATIONSHIP BETWEEN MECHANICAL PROPERTIES AND PHYSICAL DIMENSIONS OF HAND FORGINGS

The reports discussed above are prepared, in general, with the help of electronic data processing equipment. The individual test data are punched into cards with varying amounts of identifying and manufacturing data, nominal dimensions of the items of product from which the test specimen was taken, location of specimen, direction of testing, etc., so that source detail information is readily available for investigations of the effects of these variables as warranted.

These detail data cards are finding increasing use in correlation studies that provide a better understanding of the sources of variability experienced in mechanical properties. One such analysis has recently been made on the effect of dimensions of mechanical properties of hand forgings.

Hand Forgings, for the most part, are rectangular blocks of aluminum alloy produced by forging between flat dies into the desired dimensions. They are usually used where a forged part is desired but the application does not justify the expense of cavity dies. The desired contour is subsequently obtained by machining of the hand forgings.

In the early days of hand forging production, mechanical property specifications were established on limited data. It was thought that the properties were influenced by mass effect, and as a result, specifications evolved which recognized different property guarantees depending on cross sectional area of the product. As production and accumulation of data progressed, and a greater variety of sizes were made, those cognizant began to realize that there might be a different dimensional classification of these products that better described the variation in properties experienced than did the area. As a result, a program was established to study relationships between properties and dimensional characteristics of hand forgings. The accumulated detail punch cards from two forging plants, located in Cleveland, Ohio and Vernon, California, provided the necessary data for these studies.

As is the general practice in studies of process capability, the same rules mentioned earlier to assure balanced random representation of the process were applied for the data used in these analyses of mechanical properties of aluminum alloy hand forgings. To obtain a feel of the data, a number of simple scatter diagrams were first prepared in an effort to reveal evidence of strong simple relationships.

After reviewing the data by means of these simple scatter diagrams relating mechanical properties to different dimensions and combinations of them, it was decided to proceed with a multiple correlation study on the 7079-T6 alloy hand forging data using the following model:

$$\text{Ln Prop} = a + b_1 \text{ Ln Thickness} + b_2 \text{ Ln Area} + b_3 (\text{Ln Thickness} \times \text{Ln Area}) + \epsilon$$

This analysis, in general, revealed that thickness consistently correlated better with properties than did area, and that in most cases, the explanation of the variability in properties was not improved by the multiple relationship. This is shown by the example tabulation in Table III. The results of this initial investigation opened up a more extensive study of the relationship between mechanical properties of hand forgings extending over other alloys and considering other models that potentially might result in a revision of the basis on which the guaranteed properties of these products are established. This and other multiple correlation studies have been useful in clarifying mechanical property relationships.

TABLE III

7079-T6 Hand Forged Squares & Rectangles
Long Transverse Direction
Summary of Multiple Regression Analysis

	Tensile Ultimate Strength	Tensile Yield Strength	Elongation
Coefficient of Determination, %			
Multiple Regression	39.3	45.9	3.2
Simple Regression			
Ln Thickness	39.1	45.2	0.8
Ln Area	4.7	8.8	0.0
Ln T x Ln A	0.1	1.6	2.1
Partial Correlation Coefficient			
Ln Thickness	-.603	-.637	-.080
Ln Area	-.023	-.107	.083
Ln T x Ln A	-.071	-.064	.150

Note: T = Thickness, inch
A = Area, square inch

EFFECT OF SPECIMEN LOCATION ON MECHANICAL PROPERTIES OF HAND FORGINGS

It was mentioned earlier that the location of the test specimen which is removed from the product for determination of mechanical properties is sometimes of practical importance. A statistical study designed to evaluate the effect of specimen location on mechanical properties of hand forgings was undertaken to determine the necessity for standardization of this factor in applicable specifications. It was decided to conduct this investigation in accordance with a sequential design to be sure to pick up a difference in tensile ultimate or tensile yield strength due to specimen location of the magnitude of 1500 psi or larger if such existed. The design agreed upon was based on an alpha risk (the risk of concluding there was a significant difference when, in reality, there was not) of .05 and a beta risk (the risk of failing to obtain a decision of a significant difference when there did exist, in reality, a difference as great as 1500 psi) of .05. Having some feel of the expected standard deviation of the proper-

ties, charts (of which Figure 2 is an example) were calculated to give the above power to the investigation.

It will be noted that Figure 2 consists of a set of limit lines for the cumulative difference between tests made at any two selected locations. If the plot of cumulative difference in the paired test data crosses either of the outside slanting lines, we have significant evidence that a real difference exists in properties of specimens removed from the respective locations. Similarly, if the plot crosses into the angle enclosed by the inside pair of slanting lines, we have evidence that a real difference of the magnitude of the selected 1500 psi most likely does not exist between the selected test locations. The procedure followed in calculating the limit lines of the charts was that given by Davies.(1)

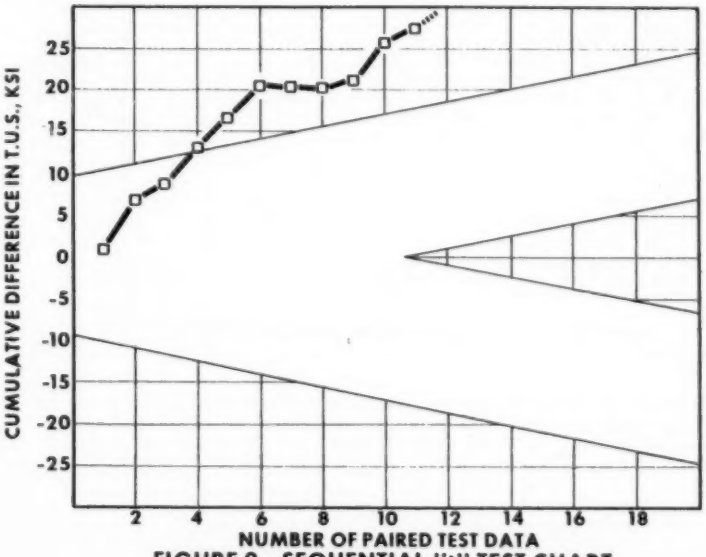


FIGURE 2 - SEQUENTIAL "t" TEST CHART

We were interested in this investigation in covering the range of sizes of forgings covered by current specifications, and therefore, certain lots of production forgings were selected as they were being planned to fairly well cover these ranges. One forging from each of the designated lots was tested in the long transverse direction with the specimens cut from the locations shown in Figure 3. The slice indicated on the right side of this figure was removed from the forging before heat treatment and then heat treated with the forging in the same heat treat load. All the specimens were cut from the designated locations after heat treatment and tested. The differences in tensile ultimate strength and in tensile yield strength for the several sets of location of test specimen were plotted on the respective charts. This testing pro-

(1) Design and Analysis of Industrial Experiments, Owen L. Davies, Hafner Publishing Co., New York, 1954.

gram was continued until decision points were arrived at on the respective charts indicating that there did or did not exist a significant difference in the properties at the given locations. In some instances, this decision point was reached after testing only a very few forgings and the testing was continued in order to establish the actual average difference within closer limits.

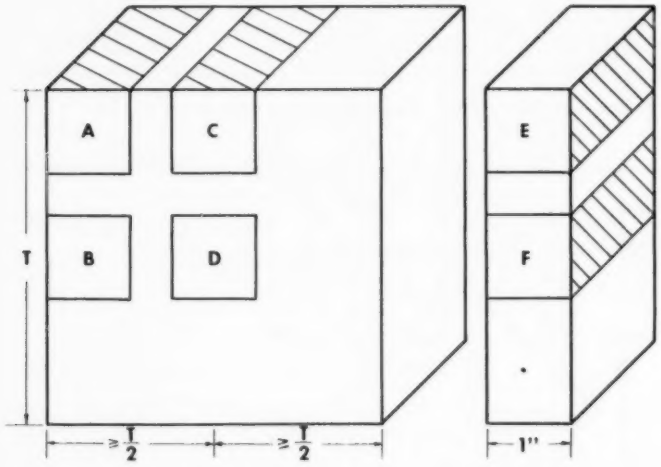


FIGURE 3 - LOCATION OF TEST SPECIMENS

Table IV summarizes the average differences in properties between several of the locations tested for one of the alloys, together with the 95% confidence limits for the average difference in each case. The locations compared in Table IV were those of

TABLE IV

Effect of Specimen Location on
Mechanical Properties of 2014-T6
Hand Forgings
Long Transverse Direction

	Specimen Location	
Property	F-D	B-D
T.U.S., KSI	2.96 ± 1.17	1.99 ± 1.18
T.Y.S., KSI	3.41 ± 1.53	1.69 ± 1.45
Elong. %	-0.77 ± 1.38	-1.37 ± 1.38
No. of Paired Data	17	17

Note: Thicknesses of Forgings tested ranged from 2 - 6 inches.

greatest interest inasmuch as the existing governing specification at the time could be interpreted to permit testing in any one of these locations for determining acceptability under the specification. Alcoa's standard test location is that identified as D in Figure 3 and Table IV.

RELATIONSHIP BETWEEN TENSILE YIELD STRENGTH AND COMPRESSIVE YIELD STRENGTH OF 7079-T651 PLATE

In recent years there has been an increasing interest on the part of customers in the compressive yield strength of aluminum alloy products. It has been Alcoa's practice for years to develop information on this characteristic of our products in our Research Laboratory. With the increasing interest in this property in certain alloys and products, it was considered desirable to extend the information available by establishing programs of compressive testing in our plants over a period of time to obtain a more extensive knowledge of the relationships between compressive yield strength and the routinely tested tensile properties.

In one of these plant program we have been studying the relationships of compressive yield strength of 7079-T651 plate with the tensile properties. Paired data have been accumulated over a period of several years by testing adjacent specimens in compression and in tension in the longitudinal, long transverse and, where thickness permitted, the short transverse directions. These data have been analyzed in several ways to develop useful relationships. In one such set of data, we had accumulated 216 pairs tested in the long transverse direction, i.e. parallel to the width of the plate. The thickness of the plates represented in the data ranged from 3 to 5 inches.

Although other relationships in these data were investigated, we were most interested in the relationship between compressive yield strength and tensile yield strength. The first analysis performed was a simple regression of compressive yield strength on tensile yield strength. The equation developed in this analysis was as follows:

$$\text{CYS, KSI} = 2.323 + 1.032 \text{ T.Y.S., KSI}$$

and a Coefficient of Determination of 0.625 developed. Figure 4 shows the regression line, 95% confidence bands and 95% prediction limits developed in this analysis.

As a matter of interest, these same data were analyzed by developing the ratios of compressive yield strength to tensile yield strength and determining the average and standard deviation of these ratios. The average and 95% confidence limits around this average were determined to be as follows:

$$\text{Avg. } \frac{\text{CYS}}{\text{TYS}} = 1.068 \pm 0.004$$

It was of interest to determine if the ratio of compressive yield strength to tensile yield strength varied with thickness. A simple regression of the ratios on thickness was, therefore, carried out and a correlation coefficient of 0.182, which was just about significant at the .05 risk level, was obtained. The regression equation developed was as follows:

$$\frac{\text{CYS}}{\text{TYS}} = 1.032 + 0.009 \text{ Thickness, inch}$$

Although the estimated regression coefficient of 0.009 is sig-

nificant, there is some doubt concerning its real magnitude in relationship to the practical importance of considering a different ratio to apply depending on thickness. Further work is being done along these same lines with similar data on other alloys which should indicate if this thickness effect is maintained and should be taken into account.

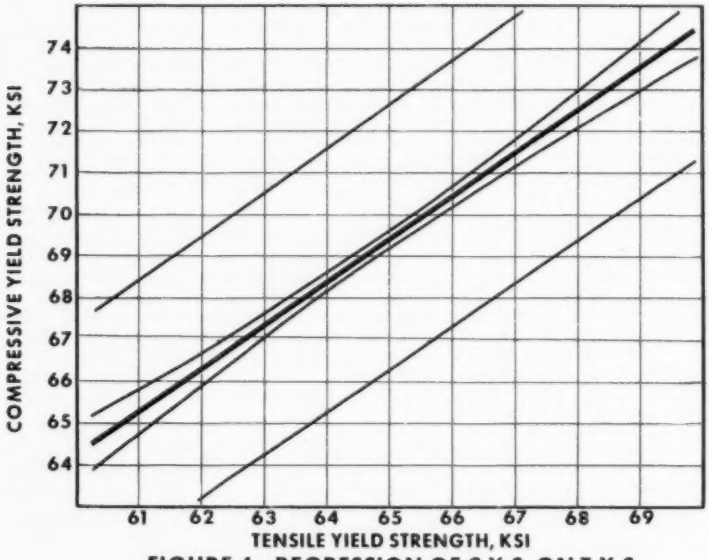


FIGURE 4 - REGRESSION OF C.Y.S. ON T.Y.S.

CONCLUSION

Variability and biases in mechanical properties do exist. By statistical studies, these can be separated and put to good use in establishing sound specification and design values. I have attempted to indicate some of the work we are doing at Alcoa directed to this end. The general program is to separate out the causes of biases from the random variation insofar as they have a practical effect on property guarantee or design values. As an example, where the relationship between properties and dimensions of a given product are reliably developed, this information is used to group sizes to the extent that different property levels among size groups permit of higher guarantees for some of the groups as compared with guarantees developed from combined data across all size groups. The extent to which size grouping is effected depends rightfully on the magnitude of difference in property level between successive size groups. In general, it does not appear of practical importance to group sizes to effect less than a thousand pounds per square inch tensile ultimate or yield strength differential between successive groups.

There is much work for statistical approaches in the field of study of mechanical properties of aluminum products and other metal products. We have done a considerable amount and will do more to develop a better understanding of this phase of our materials.

AIR FORCE QUALITY CONTROL - 1960-1962

Joe G. Schneider, Colonel, USAF, Chief, Quality Control
Headquarters, Air Materiel Command, Wright-Patterson AFB, Ohio

"Can Quality Control meet its challenge in the Space Age?" To answer the question and to realize the USAF goal of qualitative superiority, we are placing greater and continuing emphasis on the quality and reliability of weapon and support systems which are procured, stored, maintained and used by the Air Force. This emphasis poses the challenge to Quality Control - not only in the Air Force, but in Industry. The concentration of our aerospace power in fewer systems of vastly greater complexity and cost demands assurance of quality and reliability of a far higher order than we have heretofore faced. The task of realizing this assurance is made more difficult by the time-compressed development/production/test/operational cycle.

During the past several years, we have worked toward an integrated Air Force Quality Control program covering the entire life-cycle of our materiel. We are confident that during the time period 1960-1962, this program will see fruition. My purpose, in this paper, is to summarize the progress we have made to date and to chart in some detail the path ahead. This path, by the way, is not unexplored. As I shall point out later, our activities in the ballistic missile program have proved that the route, while difficult, is negotiable and its end is well worth attaining.

It is useful to view the life of Air Force materiel as a spectrum, which we might depict thus:

/ DEVELOPMENT / MANUFACTURING / OPERATIONS / SUPPLY / MAINTENANCE /

Until a few years ago, the portion of this spectrum which was significant or "visible" to quality control people-both in Industry and the Air Force - was that labeled "manufacturing". As quality control began to attain wide recognition as a function of management, it became obvious that its technique could be used profitably to protect our tremendous inventory investment.

From humble beginnings, a materiel quality program has evolved in the supply area and is now standard in all Air Force depots. The primary purpose of supply quality control is as to assure that the quality of materiel is retained through proper methods of storage, preservation, packaging, packing, and, where required, technical order compliance. Interestingly enough, one of the most important quality characteristics at this time is identification. Dependent as we are on mechanized procedures, an error of one digit in a stock number makes an item as worthless as would be a good coating of rust or a defective circuit.

In the depot maintenance area, we have a well developed quality control system which closely parallels that of an industrial activity. Our maintenance managers are as responsive to their customers' quality requirements as they are to production schedules. Naturally, the quality standards for overhauled equipment take wear into consideration and in some cases permissible tolerances are greater than those in the procurement specification. In any case, maintenance quality control assures that repair or overhaul work is accomplished in accordance with in-service engineering standards which in themselves are established in such a way as to restore as nearly as possible the original degree of quality and reliability. When depot level maintenance is performed contractually, the same maintenance engineering standards of quality are applicable. Assurance of quality is obtained through Air Force surveillance of the contractor's quality control system.

Quality Control is now a familiar term in the Air Force operating commands - SAC, TAC, ADC, MATS, etc. It is an outgrowth of the old organizational and field maintenance inspection system. Many of the techniques of modern quality control are employed to assure that, at the point of use, the quality and reliability of our weapons and supporting equipment are kept at the highest possible level.

With the implementation of quality control programs in supply, maintenance, and operations, we are concentrating on ways of improving their integration with the rest of the system. One of our most active endeavors in this respect is the Air Force-wide calibration program. Each of our Air Materiel Area depots operates a standards laboratory. We have established calibration laboratories at 150 Air Force Bases throughout the world, enabling them to provide complete calibration service for all test equipment required to support their assigned weapons. This system, in conjunction with contractors' calibration systems, assures the Air Force that the accuracy of measurements is traceable to the National Bureau of Standards throughout the entire life cycle of a weapon or item of equipment.

Supply, Maintenance, Operational, and Procurement elements of the Air Force are loosely linked, quality-wise, with various reporting systems for failures and discrepancies. This is a link which we are endeavoring to strengthen by integrating quality control needs into an Air Force-wide automated logistic data system.

Thus, with quality control programs established and functioning in supply, maintenance, and operations, the right hand side of our spectrum is suitably covered. Let us then turn our attention to the left hand side.

For many years, AF Quality Control has had some activity during the development phases of procurements - particularly of aircraft, missiles, and major equipment items. As a minimum, we were concerned with the acceptance of certain experimental items and the witnessing of qualification tests. In the early days of guided missiles, the beginnings of concurrent development and production brought about a revision to the Air Force Quality Control Specification MIL-Q-5923, whereby modifications of certain quality control system requirements were permitted during the early stages of a program. This version of the specification clearly established the legitimate interest of Air Force Quality Control during development. By contract requirement, it expanded the scope of industrial quality control into this area as well. The current DOD Quality Control Specification, MIL-Q-9858, does not recognize the distinction between development and production as far as requirements for a quality control system are concerned. However, Air Force administration of this specification does provide for appropriate application of the various requirements as a weapon system develops through its successive phases. I will now leave the area of "conventional" quality control, as it has been practiced by the Air Force-Industry Team for many years - leave the security of the straight and narrow, well-paved road and begin to chart the path ahead.

Some two years ago, we recognized the necessity for closer alignment of quality control activity with weapon system management and the need for greater emphasis on defect prevention and the minimizing of logistic support problems. The staff Quality Control elements at Hq Air Materiel Command and the Ballistic Missiles Center developed the "Air Force Quality Control Program for Ballistic Missiles and Space Systems". Because of the success of this concept, it is being made applicable, with appropriate refinements, to all weapon and support system procurements. It is being extended, initially, to selected systems. Clearly, this approach to missile quality control has not solved all missile problems - not even all the missile quality control problems. However, we are satisfied that this is the approach we must take - its success depends as much on the Industry partner of the team as on the Air Force. I will amplify this point later.

Before discussing the operating concept of the "Standard AF Quality Control Program for Weapon System Procurement", the following comments should be considered to assist in establishing the proper reference frame.

> "Quality Control", as used in this paper is less adequate than some such term as, perhaps, "quality assurance". It is used in full recognition of the fact that, as an organizational element, it does not exercise control over many of the elements with which it is, of necessity concerned.
>
> The expression "quality and reliability" is used frequently. This is because reliability is widely recognized as a major problem area and because we have no universally accepted statement of the relationship between quality and reliability. The Air Force views quality as Effectiveness, which is a function of a number of "quality criteria", including reliability, performance, maintainability, and accuracy. Hence, Reliability is a part of quality and Reliability Assurance is, in part, a responsibility of Quality Control.
>
> Two predominant characteristics of weapon system procurement are time-compression and contractor-development of engineering data. Both of these characteristics aggravate the problem of timely approval by the Air Force of firm engineering data. Since such data are the basis for both acceptance of hardware and logistic support, the Air Force must minimize this problem.
>
> Every advance in weapons and military equipment has demanded an increase in the average technical competence of every group involved in their development, manufacture, care, and use. Quality Control is no exception. As far as the Air Force is concerned, the function of assuring compliance with contract requirements - a necessary condition for acceptance of hardware - has become such a highly technical job that we have found it necessary to augment the staff of the AF Quality Control Representative (AFQCR) with appropriately qualified engineers. Increased complexity, reliability demonstration requirements, and less specific contractual definition of acceptability are among the foremost reasons for this. We now have more than 150 quality control engineers associated primarily with ballistic missiles. This admittedly only scratches the surface - but the need is clear-cut and the benefits have been amply demonstrated.

Our "Standard AF Quality Control Program for Weapon System Procurement", while perhaps having revolutionary aspects when compared with the World War II inspection program, is in reality but a logical evolution. It involves no change in organizational authority or responsibility. In particular - and this should be kept in mind throughout the following discussion - it involves no responsibility for engineering approval. This program is oriented towards:

> Facilitating the Air Force's legal requirement of assuring the compliance of materiel to contract requirements - in other words, determining the degree of attainment of acceptance criteria for hardware which in many cases is being procured on the basis of performance specifications.

Emphasizing early prevention of quality problems, largely through timely feedback of pertinent quality information to the appropriate Air Force decision making authority - usually the Weapon System Project Office (WSPO).

Let us turn our attention to specific features of the program. Attention is focused on development activity, since preventive efforts at this time offer the greatest opportunities for pay-off. Air Force Management, embodied in the WSPO, is constantly making decisions which affect the overall quality of the weapon system. Such decisions are related to reliability program requirements and results; test program approvals; acquisition and approval of engineering data; calibration; management review points such as mock-ups, development engineering inspections, and contract technical compliance inspections; and the all-important area of Engineering Change Proposals (ECP) and waivers which may have an adverse effect on quality. Better decisions, quality-wise, result when the WSPO can take advantage of the day-to-day project knowledge of the Air Force Quality Control Representative (AFQCR) at the plant. The Quality Control Engineer plays a major role because of his technical competency and his ability to communicate with both contractor and AF engineering personnel.

It will be obvious during the following discussion that one or two AF quality control engineers at a plant, working unilaterally, can only scratch the surface. As with other quality control functions, those of the quality control engineer must be performed on a surveillance basis.

Reliability Program - Most new Weapon System contracts include the requirement for a contractor reliability program, in accordance with either MIL-R-26674 or AFBM 58-10. WSPO management of this program is facilitated by on-the-spot reports from the AFQCR who conducts plant level surveillance over the implementation of the program. Such reports are submitted either on his own initiative or at the request of the WSPO when a particular area is indicated for special attention.

Contractor-Prepared Product Specifications - Since the weapon system contract initially specifies broad performance requirements, the progressively more-detailed product specifications are developed by the contractor. The AFQCR, largely through his quality control engineer, keeps abreast of the development of such specifications and makes in-process reviews of the aspects which are of concern to Quality Control. He is particularly interested in determining that appropriate inspection and test requirements are provided, that specified tests are realistic and will actually demonstrate the achievement of performance parameters, and that specified tests are compatible with available test equipment. He is also concerned with the compatibility for related systems, subsystems, and components. An obvious, but all too frequently violated principle, is that the specifications must provide a valid description of the developed hardware. This, together with timely submission of the specification for AF engineering approval, if required, is essential if we are to have any valid basis for hardware acceptance and for logistic support actions.

Monitoring Review Points - Certain benchmarks for comprehensive review and evaluation of development progress, such as Mock-up inspections, Development Engineering Inspections, and Contract Technical Compliance Inspections have long been standard tools of AF program management. The AFQCR has in the past participated to some extent in the first two and Quality Control at Hq AMC, in the third. To take full advantage of the on-the-spot knowledge of the AFQCR and his staff, their participation in all three has been increased. The same is true of the recently established Reliability Review points. Greater emphasis on Quality Control inputs to these activities results in detection and identification of quality and reliability problem areas early in the program.

Classification of Characteristics - This is a preventive technique being introduced into weapon system contracts. Characteristics are the attributes of a part or an assembly. Classifications are the design criteria establishing importance of the characteristic. Tolerances are the design criteria establishing permissible variation of the characteristic. The classifications are an aid in producing and maintaining quality and reliability by providing a common basis for placing emphasis on the more important characteristics during all phases of production, maintenance and operation. Characteristics of a lesser importance may be produced, inspected, maintained, etc., by means requiring less emphasis.

Failure Data Feedback Systems - Contractor failure data feedback systems are required to begin early in contract life. Such systems must include provisions to assure that effective corrective actions are taken on a timely basis to reduce or prevent repetition of the failures. AFQCR conducts surveillance over the administration of these systems and verifies the appropriateness of technical analysis and actions, based on the failure data. It is hoped the contractor's system will permit transfer of the data into the AF system with minimum effort.

Team Acceptance - Time compressed development and production programs dictate the need for active quality control participation in weapon system programs at their outset. "Team acceptance" is a procedure which provides for joint ARDC-AFQCR review of contractor furnished data approximately 30 days prior to final functional test of each major item to establish understanding of the minimum acceptance criteria for the end item. These preliminary data are used as basis for engineering approval by the ARDC members of the team and then used as the acceptance criteria by the AFQCR members. Such team action of concurrent approval and acceptance facilitates time compression objectives, supports the WSPO during the early phases of the program and aids in developing AFQCR's technical competence.

Statistical Methods - Contractors are required to incorporate the use of statistical methods wherever meaningful and practical throughout all phases of the weapon system program. This includes the application of such methods as design of experiment in the development phase, as well as the use of process control and acceptance sampling techniques in manufacturing. Plans for demonstrating achieved reliability are required to include specification of risk of confidence level. Proper use of appropriate statistical methods will result in greater assurance and reduced costs. In destructive testing, their efficient use is critical. It is important that methods to be used are considered early in planning to achieve maximum efficiency. The AFQCR reviews various statistical methods and procedures for appropriateness and validity, and verifies their proper application.

This discussion, then, portrays the essence of the "Standard AF Quality Control Program for Weapon System Procurement". It clearly demands the use of engineering skills. It neither involves nor implies that Quality Control exercises any engineering design or approval function. It does extend quality control responsibilities, as contrasted to mere interest or concern, into the early stages of development. The purpose of this is to identify potential quality problems as early as possible and to feed back appropriate information to the AF decision-making authority.

As this program, successfully used with several of our ballistic missile contracts, is extended on a selective basis to all weapon and support systems, we will have provided quality control coverage of the spectrum of Air Force materiel. We will have a completely integrated quality control program, from development and production through the repetitive cycle of operations, maintenance and supply.

Just as the Air Force is vitally concerned with Quality Control in areas other than manufacturing, Industry's role under the Weapon System Concept is placing pressure on industrial quality control to expand its horizons.

In many cases, the prime weapon system contractor and many associate and subcontractors as well, have responsibilities for maintenance and supply even after operational deployment of the weapon system. As requirements for contractor technical support of the operating commands increase, so will the requirement for quality control communication. Active quality control programs in Air Force supply, maintenance, and operational activities provide coverage on our end of the communication link.

The question posed at the beginning of this paper can be answered in the affirmative if the Air Force and Industry continue to cooperate in the same dynamic and far-sighted manner that characterized the evolution of the surveillance program, which is now embodied in Specification MIL-Q-9858. In particular, we say that Air Force Quality Control activity during development employs the techniques of surveillance. This is of course meaningless unless the contractor's system lends itself to surveillance.

In the final analysis, the weapon system and associate contractors must design, test, produce, and deliver the hardware. Consequently, it is the contractor who must meet the challenge of achieving specified quality and reliability requirements with adequate management programs. Weapon system specifications establish the general requirements for, and require the contractor to implement these programs. The reliability program requirement (MIL-R-26674 or AFBM Exhibit 58-10) and the quality control system requirement (MIL-Q-9858) are similar in concept and complementary in practice. Both require a program throughout contract life. Plant level assurance that these programs meet contractual requirements is a responsibility of AF Quality Control. The only difference in the AF management function is that responsibility for approval of the Contractor's Quality Control system rests with the AFQCR whereas responsibility for approval of the Contractors' Reliability Program rests with the WSPO. Consequently, surveillance over the contractor's quality and reliability efforts are so closely related within the industrial complex as to be virtually inseparable.

The challenge to Industry, as I see it, is to follow the cue of certain leaders in the field by achieving effective integration of the quality assurance function throughout the life cycle of the contract. Only in this way can we really have assurance that our weapon systems and supporting equipment do, in fact, meet the requirements of the contract. Quality Control can thus fulfill its role in supporting the USAF goal of qualitative superiority.

ARMY PROGRESS IN QUALITY CONTROL, CALENDAR YEAR 1959

Colonel Eldon A. Koerner, USA, Chief, Standards Branch
Office of the Deputy Chief of Staff for Logistics, Department of the Army

For the past several years it has been traditional for the Military Departments to tell you, during your national convention, of progress made in Quality Control and to give you an advance look at areas of Quality Control we intend to explore.

One of the major revolutions in the Industry-Army relationship was initiated in 1957. It was at this time that we sharpened the distinction between the responsibility of Industry in Quality Control and that of Army. Since the Air Force, up until a few years ago, had by far the most complex weapons, and I am thinking now of the airplane, it was far in the lead in realizing that the control of quality of complex weapons could not be handled on an after-the-fact of production basis. Their specification MIL-Q-5923 was a trail blazer in this regard. All of us in the military system now understand the need for such a specification and, as you know, a coordinated military specification MIL-Q-9858, was agreed upon and published over a year ago. This specification, coupled with an Army Standard (MIL STD 643) spells out in general terms the responsibility of Industry and of the Army.

What was not so well understood by many of us up until a few years ago, was how to establish clearly the division of responsibility between Industry and Army on simpler items, those which can be and are covered by definitive specifications which describe the item and the inspection or quality assurance procedures necessary for assuring that adequate quality has been produced.

For years standard contracts have required that a contractor maintain an acceptable inspection system. Of course a requirement for a system without spelling out what an acceptable inspection system consists of tends to be ambiguous. It was difficult, almost impossible, to come to a meeting of minds on when a contractor actually had an acceptable inspection system.

In 1957 Army announced that in the future a definition of an acceptable inspection system would be as follows:

> A contractor's inspection system is acceptable when it includes, as a minimum, the inspection, including testing operations spelled out in the specifications or supplement thereto; and
>
> When the produced quality proves to be adequate when verified by the Army.

In the past three years this last policy has lead to serious misunderstanding as to just what the responsibilities of the Department of the Army were when Industry was responsible for performing the inspection operations spelled out in the specification. This misunderstanding has run the gamut. Some people believing that the Army had given up all of its prerogatives in inspection, even up to and including the authority to make the acceptance decision.

I wish to highlight at this point the fact that the acceptance decision is always made by a competent United States official. The authority for such decisions is not delegated to Industry. Industry responsibility is clearly that of producing the items and developing evidence regarding their quality. Army will audit Industry activities both in the planning of the inspection programs and in the operation of the programs. Army will also perform such product inspection as is found necessary to verify that the programs are doing what they ought to do. On this basis, decisions to accept or not to accept are made.

Just recently, in attempting to coordinate specifications which spell out Industry responsibility for inspection, Army became aware of a close analogy between the documents needed for the complex weapons systems (MIL-Q-9858 and MIL-STD-643) and the documents needed for simpler items. We are now in the process of developing a general specification to establish the responsibilities of contractors producing such items. This specification will do for simpler items what the two documents for complex items do. Section III of this specification will describe the general responsibility of the contractors and Section IV will establish the basis that Army will use for making its acceptance decisions. Of course the complete package will include the detailed item specification for specific inspection operations.

A few minutes ago when I started I outlined this talk as being in two parts. First, progress made, and Second, a look at future actions. I see that I have, in one sense, put some of the future actions in before the progress.

Regarding Progress. By the end of calendar year 1958 Army had, through use of Quality Control specifications similar to MIL-Q-9858, imposed mandatory requirements for quality control on about 75 producers of complex items. More important, Army had been able in these plants to significantly reduce the detailed operations performed by Army personnel because Industry was clearly producing a high quality product. Larger savings were made in the area of simple items. By the end of 1958, Industry had been able to develop adequate contractor inspection and quality control operations on a voluntary basis in about 650 plants. Estimated savings through calendar year 1958 in salaries alone amounted to about $1 million.

I am proud to say that this progress through calendar year 1958 has continued at about the same pace through 1959. Adequate quality control required by use of MIL-Q-9858 type specifications now is in operation in about twice as many plants as in 1958. The same situation holds true for voluntary quality control of simple items. There are adequate plans in over 1300 plants. Industry and Army have been able to agree upon the adequacy of their operations with further savings to the country. We look forward to continuation of this progress through this year. We believe that the specification describing the general responsibility of contractors in the inspection and quality control of simple items will further enhance this program.

Looking forward, Army intends to investigate and, if possible, develop better understandings and policies regarding several difficult areas in the relationships of quality control to the rest of the designing, producing, using and managing functions of Army and Industry. These include calibration problems, reliability problems, the use of Warranty systems, and a study of responsibility both in Industry organizations and within the Army regarding quality of design and quality of conformity to the design.

We have a lot to learn about calibration. Not that we haven't, in the past, had calibration programs adequate for our needs at the time. Our needs are changing rapidly.

Calibration of tools and instruments has always been necessary at the time of production, at the time of quality control, at the time of acceptance, and during maintenance and use of the materiel produced. For these calibrations to have meaning, they have had to go back to a common set of standards, those maintained by the United States Bureau of Standards. Back in the days when safety factors were designed to be in the neighborhood of five or ten or more to one, calibration was not nearly so critical. Today, with our complex weapons containing literally thousands and thousands of components, all of which must operate with split second accuracy, calibration has become more critical than some of us realize. Aggravating this situation is the number of instruments and tools that must be calibrated. This calibration is made extremely difficult by a chain of ever decreasing tolerance as we approach the basic standards maintained by the Bureau of Standards. For example, if a production part has a tolerance of a thousandth of an inch, we normally try to gage this with a gage that has a tolerance somewhere in the neighborhood of a tenth of a thousandth. This is not too difficult, but when we take the gage and attempt to check it, we need laboratory instruments capable of measuring and giving accurate results within a tolerance somewhere in the neighborhood of one hundredth of a thousandth. Of course, this laboratory instrument must be checked for accuracy and this leads us to a requirement for accuracy down in the neighborhood of a few millionths of an inch. Fantastic as it may seem, we still have one more link in this chain of circumstances. This leads us to master gage blocks and comparisons by interferometer, thus achieving an accuracy somewhere in the neighborhood of one millionth of an inch. Now as long as only a few of us expected the National Bureau of Standards to provide this chain of calibration, the Bureau could accommodate us.

Recently, the realization that vitally needed calibration was not being accomplished adequately resulted in the Bureau of Standards receiving more and more calls for calibration at all levels in the chain I've just described. This is not the type of calibration with which the Bureau of Standards should be burdened. Essentially, they should have to carry the load only of occasional verification of primary standards maintained by the user so that he may calibrate secondary standards down through the chain to his own tools and instruments. Because of the unwarranted burden placed upon the Bureau of Standards it currently is having to "back order" calibration requests unless they have the highest priority.

The Military Departments have a sizeable calibration program, including their own primary standards, which relieves the Bureau of Standards of many of the details of lower order calibration operations. I believe Industry is going to have to have a similar program, including primary standards, available throughout the country to its members, so that the load on the Bureau of Standards will be reduced to an even more realistic demand.

What I have said about calibration in terms of sizes, such as length, also applies in a more troublesome area - that of electronics. When you consider that some of our electronic switch designs call for cycling times down in the neighborhood of a millionth of a second, you realize that calibration in the electronics field also requires fantastically accurate comparisons. One example of the fantastic accuracies required is in Army Ordnance primary standards which are considered to be accurate to one part in ten million.

As our demands for accuracy become more and more stringent our ability to maintain the ten to one ratio in each step of the calibration system probably will be something we cannot achieve. In some instances calibration systems now are operating with the increase in accuracy in each step being less than ten times the accuracy of the previous step. This is going to require study.

Another problem in calibration sometimes lost sight of is the need for an integrated calibration program covering all facets of measurement from the cradle to the grave of an item. There is no point to adjusting or checking a system at the time of design or user test with instruments that are of a different order of accuracy than those that will be used when it is produced, or accepted, or maintained. The same orders of accuracy must be available throughout the life cycle of an item.

One of the elements of Quality Control that is most difficult to get a grip on is its relation to reliability. All of us realize that, if quality as it relates to conformance with design is poor, reliability is bound to suffer. Another and more difficult part of the reliability quality control problem has two aspects: The first is, how do you specify reliability as a meaningful requirement, and, how do you prescribe numerical tolerances for it. Even more difficult in this regard is how do you make meaningful measurement if you succeed in prescribing these numerical tolerances. Since reliability is the essence of the item --"Will it work or will it fail" -- we must have reliability. We must have a high degree of assurance that the reliability we have specified has truly been achieved. Since true reliability is a guarantee that the item will work, sometimes for a considerable period of time, the tests for it are almost always expensive. For items such as missiles that have a one shot application the tests are both destructive and expensive. For other types of military equipment we want long, trouble free life. Testing for this capability takes a long time; is expensive; and, in many cases, while not destructive, changes the item in such a fashion that we don't dare use it thereafter for military operations. The items you gentlemen make for us take a long time to produce. They are continuously undergoing change for improvement. Even if we had more money to spend, under our present concepts we may never have enough of these items to make a sample large enough to give the statistical assurance of reliability that we would like to have. The second aspect of the reliability characteristic in quality control, the quality of design, is even more difficult to specify and measure. In missiles and, in fact, in most of our complex weapons, our contracts call for designing and then producing. How do you inspect, how do you measure, the quality of a design? How must we organize to accomplish these two tasks? Quite simply, you specify the quality of design required when you establish the task you want the weapon to accomplish. After it is designed and a number are produced, you can check to see if the weapon accomplishes the specified task. But there are high expenses and other troubles involved in taking large samples to check design quality. If we abandon this approach, we must attempt to measure quality of design at the time the design is first developed, before production has started.

People often fail to realize that when we contract for design and production of a new weapon we have a two part contract. One is for services, the other is for materiel. Military Quality Control has always by definition claimed to assure the quality of both services and materiel. When the service is one of designing, it is almost impossible to measure its quality, and, therefore, its acceptability, without duplication of the engineering effort that was required to first develop the design.

It may be that this last thought, that of duplication of the engineering effort, is not as bad as it first seems. For years in critical and complex items we have used the idea of redundancy to reduce the likelihood of failure. Possibly redundancy in our system of designing these complex items may prove to be one of the better ways to increase reliability. Another thought that comes to mind on this question of reliability is the aversion people have to the high cost of designing, development, and testing of an item as related to the cost of the items that will be produced once this is finished. In years gone by it was common to measure the efficiency of a program in terms of cost of development as related to the mass-produced cost of the item. This concept may be inadequate when we think of things such as ballistic missiles and other complex weapons systems. It may be that the most economical system for getting these in useable form into the hands of our military using activities may be one of high development and design cost with a relatively lower cost of producing the required number of items.

The last facet of Quality Control I wish to discuss today is that of Warranties or Guarantees. You may ask "How does this relate to Quality Control"? My answer is: If we had a clean cut well working guarantee or warranty system which assured timely replacement of nonfunctioning items that were the responsibility of suppliers we would be able to accomplish the mission of having acceptable supplies in the hands of the users without any need for an Army inspection or Quality Control system as we now know it.

In discussing warranties or guarantees several things come to mind. First, many of the items the Army buys are covered by a standard commercial guarantee normally given by the trade and not the result of any special extra costs. I realize that such guarantees, if they are to do any good, must cost somebody something, and that the cost is an inherent part of the cost of an item. However, these costs normally are not broken out by the industry concerned. Even if we refused to accept such a guarantee there would probably be no discount allowed. Army policy regarding this type of guarantee is to accept it without question. In effect, it costs us nothing that could be saved. One problem we do have with this type of guarantee is frequent failure to use its provisions when something fails in our system. When an item is out of order for reasons covered by the guarantee there is a strong tendency to call in one of our own technicians and get it fixed as soon as possible. This, of course, nullifies any guarantee. Another problem with this type of guarantee exists due to our practice of buying and storing supplies sometimes for considerable periods before using them. The guarantee period expires before we start to use this item.

There may be a way to help this situation. In commercial practice a producer often encloses in the packing or shipping container a guarantee card to be filled out at time of purchase by the consumer. This time may be a number of months after production and sale to the jobber or wholesaler. The guarantee period commences at the time the consumer takes possession.

We are exploring the possibility of applying this concept to Army supplies. The contractor, under this concept would agree to a guarantee period that would start at the time we take the item from stock and place it in use. Such a practice would, of course, be quite analogus to commercial practice.

Another type of guarantee system is used by the Army in cases where there is a sole source supplier. Some concerns will sell only under special guarantees. There may be clear and stated costs for such guarantees, but to get the needed supplies we have to accept the guarantee and pay for it.

These two types of guarantee are not troublesome to us. The last and most troublesome part of the guarantees that may be used is that where we develop special Army clauses for insertion in contracts. The costs of these guarantees are difficult to assess. In the few instances where we have been able to estimate the true cost of such guarantees it appears that somewhere in the neighborhood of five percent of the item value is charged. When we contrast this five percent figure with the cost of what we now consider to be an adequate Army Inspection and Quality Control program, normally one percent of the item value, we are given cause to think that possibly special guarantees imposed by the Army are not an economical practice. Currently, my office is making a study of costs of guarantees and ways of using them and will develop an overall Army policy some time this year.

NAVY PROGRESS IN QUALITY CONTROL FOR 1959

Captain G. L. Heap, USN
Assistant Chief of Naval Material (Field Services)
Office of Naval Material
Department of the Navy

INTRODUCTION

A few days ago, I came from the East Coast to San Francisco in about five and one-half hours. Such speed is commonplace. Yet when I entered the Navy, the same trip took five and one-half days. At that time, today's speeds were not predicted even in such literature as "Tom Swift and His Giant Aircraft." The scientific advances of the past three decades have been due to many things not the least of which is "Quality Control." My remarks today on the Navy Progress in Quality Control during 1959 concern the development of quality assurance as related to the production of ships, aircraft, weapons, and supplies for your fleet. I shall also speak of progress in improving the technical calibre of quality assurance activities. These subjects are intimately related. Progress in both areas has enhanced the economic procurement of a wide variety of reliable fleet items.

DEVELOPMENT OF QUALITY ASSURANCE

The principal areas of progress in the development of quality assurance have involved refinement of our inspection methods in response to the technological improvement in the instruments of inspection. The marriage of nuclear propulsion and ballistic missiles in submarines now in commission has posed production problems in unbelievable complexity. In response to the urgency for production of such vessels, quality assurance techniques sensitive to the new technologies, but utilizing the existing organizations and procedures was mandatory.

This involved the application of technical refinement where needed, simultaneously with improvement in less difficult and less expensive methods where suitable. In addition to the more exotic newer weapons, our quality assurance procedures must be responsive to the requirements of a myriad of less sophisticated ships, aircraft, weapons and supplies. The application of quality assurance principles differs with the product and with the producer. Aircraft, ships, precision machinery, nuclear power plants, electronics and ordinary consumer items each require different standards and different technologies. The Navy's inspection offices must be prepared to exercise judgment in the selection and application of specific techniques most appropriate to the product and the industry involved.

Items such as nuclear components, missile guidance systems, and ships inertial navigation systems all involve quality characteristics which have minuscule tolerances that are interrelated and that are quite difficult to verify. These products are characterized by rapid design obsolescence. Their development and production stages have been compressed into one continuous process. Quality assurance is involved intimately both in the fabrication and in the control of fabrication processes. There must be a realistic development of supplier responsibility for every phase of quality, despite the difficulties normally associated with specifications and the documentation for such products.

Quality assurance on these complex items involves the evaluation of the suppliers fabrication methods and his control of each stage of production. These evaluations must be conducted by survey teams of technical competence far greater than that required only a few years ago. To fulfill this requirement, we have raised the qualifying requirements and adjusted the grade levels in Navy inspection offices, particularly in the engineering specialties.

This has not been easy since industry has been searching for these same types of personnel. However, we have been reasonably successful during the past year, and I feel that we meet 1960 well staffed. The responsibility for quality control surveys is placed at the field office level. This permits the supplier to receive instructions from, and to be monitored by, only one office. When certain new or exceptionally complex items are involved, we augment the survey team with specialists from the appropriate technical bureau. In the case of nuclear components, we may use personnel from prime nuclear power plant contractors. With such a team, we feel competent to evaluate the quality control system of any supplier. It is appreciated that knowledge of non-destructive test methods, sufficiently objective to provide positive assurance, is limited particularly in radiography and ultra-sonic testing. Employment of these methods is somewhat hampered by a shortage of competent personnel, but progress is being made and evaluation surveys are not limited to complex items. For all items, partial surveys of supplier's inspection methods and equipment have been conducted to properly utilize the supplier's own objective quality data, and to adjust government verification inspections to the points of greatest need.

Application of specification MIL-Q-9858, which began on a voluntary basis in some industries, was not as extensive as desired in 1959. This was due to the late publication of the specification, and to the time involved in implementing it in contractual documents. We are applying this specification now to shipbuilding, nuclear components, and missiles. The standard clause for supplier responsibility to perform all inspections and test--except those specified for government performance--is still not spelled out in all specifications. Quality assurance techniques have been applied by voluntary agreement with suppliers who recognized the benefits to be derived. This is particularly true where surveys have pointed out method deficiencies not heretofore recognized.

For a contractual backup on quality assurance, and for a clearer understanding by contractors and vendors of their obligations in connection with source inspection, there was recently published the "Manual for Source Inspection and Administration of Navy Procurement." This is available at all inspection offices and is essentially the same as the general specification for inspection of material, which it supersedes. It utilizes current armed services procurement regulation terminology and condenses, clarifies, and modernizes the principals previously set forth in the general specification. It is consistent with specification MIL-Q-9858 for contractors quality control systems. It is intended that the latter be applied to the procurement of complex items. The manual will be applicable to all navy procurement requiring inspection or contract administration at source.

Improvement of the actual quality of our ships, weapons and supplies requires quality assurance action at many levels. This action must take place in the design, manufacture, and source inspection, and it must continue through the installation, repair, overhaul, and user stages. It is not enough that a product performs satisfactorily at the manufacturer's plant, it must perform satisfactorily after shipping and installation. It must continue to perform at sea where there are no manufacturers' specialists available.

Too many times equipment deficiencies have not been reported back through the complete chain for information of designers, specification writers and others who need to know. We are vigorously pursuing a quality feedback system, whereby all complaints are reported to the quality assurance activities at the manufacturing level. Our field inspection offices are taking action to remove the basic cause of the trouble. They advise the suppliers of corrective actions required. They advise the contracting officers when contractual remedies are applicable, and they advise the technical bureaus responsible for specification or design improvement. Trouble reports are thoroughly analyzed to determine basic causes and to prevent repetition. Scientific methods are applied to improve our over-all quality position within budget limitations.

IMPROVING TECHNICAL CALIBRE

These developments also involved the improvement of the technical calibre of our quality assurance activities. All government organizations operate under manpower and dollar limitations. The Navy inspection offices are no exception. However, response to progress is not being retarded as we use all available means to improve the competency of our staff through both recruitment and training.

The increased complexity of products required increased numbers of professional engineers in inspection offices. This included not only the specialized offices handling shipbuilding and missile work, but also the offices which handled many of the components of these complex items. These engineers functioned not only as staff personnel, but as their numbers have increased, they were given increased responsibilities to act on technical deviations and to review and approve engineering drawings. This delegation of authority is essential in cases where production must be expedited, and where technical obsolescence is rapid. Some quality assurance documentation is generated by field activities, particularly in missiles and nuclear work. On-site technical review is essential to obtain realistic documentation which agrees with the hardware actually being shipped. Further improvements in this area are being studied.

Organizational changes, begun in 1959, are still in progress. These involve the inclusion of more technical personnel in the general inspection offices. These personnel are responsive to the technical bureaus for engineering action. We avoid setting up special offices since the existing ones can be staffed to do the job.

The training of quality assurance personnel has involved both the principals of quality assurance and the specific technical skills necessary to inspect new products. The Bureau of Naval Weapons is now conducting a continuing course for training middle management supervisory personnel both in quality assurance and in the evaluation of methods of control over the production of complex items. Training for nuclear component and nuclear core inspectors has been refined. A training program in new welding standards is being implemented. Currently, non-destructive test techniques are being taught to Navy inspectors by the Army's Watertown Arsenal in a special course oriented to material inspection needs. This training provides the basis for further training at our nuclear schools. A course in special optical tooling has been instituted at the Army's Frankford Arsenal. Other specific skills are being acquired on a local basis, both by in-service and by after hours training. Much credit is due to field personnel who participate extensively in these courses, all of which involve hours of study beyond the normal work day.

We have concentrated on improving the calibre of our supervisory personnel in respect to management know-how, quality assurance, and specific technical know-how.

SUMMARY

I believe that the Navy progressed rapidly in 1959 in the field of quality assurance. I believe that the new specifications and our new manual will be extremely helpful in assuring quality products. Our offices have met the challenge of improved technology. They are geared to the application of the newest techniques, both in quality control and contractual requirements. Our training programs, together with our higher calibre personnel, will cope with future developments. We are ready, willing, and able to continue our close cooperation with industry to insure that the Navy receives the best product it is possible to build.

Navy inspectors live by a creed.

I am an inspector of naval material. I serve in an organization dedicated to the acceptance of quality material in accordance with specifications delivered on time and suitable for the purposes intended.

I believe that naval material inspection must be guided and governed by the highest standards of conduct and ethics striving always for an ending in all things affecting the conduct of our efforts as would make reputation an invaluable and permanent asset.

I believe that mutual trust and respect exists between industry and government.

I believe in equal opportunity and in the human dignity of the common man.

I will make every possible effort to develop and maintain an efficient, loyal, aggressive organization.

I will insure that this organization receives the necessary training to keep it abreast of technological advances and that it has adequate facilities and tools with which to work.

I will never be satisfied with less than the best possible results of attainment in each and every problem to be solved.

I will be consistent and persistent in the application of the Navy's time proven policies.

I will be alert to new trends and developments and will incorporate them into our organization whenever efficiency or effectiveness may be improved.

My slogan is Quality and Service. Quality, the highest obtainable from American industry - Service, the best possible to be rendered to the technical bureaus and offices and to the contractors.

INTEGRATED SQC IN FOOD PROCESSING

C. R. Sheaffer, Quality Manager, Frozen Food Division
Stokely-Van Camp, Inc., Oakland, California

A completely integrated SQC program for food processing is indeed a very optimistic goal. Strict interpretation of the term "integrated" would indicate that this title implies SQC application in all the process of production, manufacturing and marketing. Certainly the scope of such a project would likely be beyond reality for most food processors. Perhaps more realistically we should consider a "balanced" SQC application. Such program would likely progress in the direction of completely integrated SQC.

As we consider SQC applications throughout each phase of the food processing operation, we can readily realize the importance of efficient use of SQC at each point. Failure or inefficiency in any phase greatly nulifies the efforts made in each other part of the overall operation. In fact, failure at any point may make either previous application or subsequent application highly impractical. The emphasis that should be placed at any point in the overall process would depend upon the proper "balance" of the particular operation.

The examples that will be described herein, represent only a few applications from a relatively extensive program, but I believe they will suffice to illustrate the scope of a "balanced" or potentially integrated SQC program for food processing. I will attempt to emphasize the need for such a balanced program by means of examples from a frozen pea operation.

The entire Stokely-Van Camp operation may be considered as vertically integrated since its food packing program is widely inclusive from seed selection, improvement and production to vegetable and fruit crop growing, purchasing, packing, warehousing, shipping and sales to consumer markets.

The seed program as a phase of the Company's divergent activities, all of which must be highly coordinated and accurately controlled to yield the expected results, itself must be well controlled in order to obtain the sought after overall general benefits for the Company. Accurate control application to the seed operation includes attention to the numerous aspects of varietal and strain testing in experimental plots in field and greenhouse; selection of superior sorts from hybrid, pedigreed populations; seed increase and production of proven lots under constant surveillance for purity of strain; seed crop milling under line quality control for seed condition, defects, germinability, adequacy of seed protectant application, correct fill weight, etc.

Varietal testing of newly developed or introduced innovations involves comparative evaluation of collected data obtained from test plots arranged in experimental designs, studied statistically by analysis of variance method. Adaptation, yield, quality, maturity, etc., all important plant characteristics, are investigated for their interrelationships and interactions through analysis of covariance. Development of new improved varieties is by means of scientific plant breeding, and this is concerned primarily with the science of genetics which is highly mathematical. Here much of the collected hereditary data are analyzed by use of the Chi-square test, where observed and theoretical values are compared.

Field sampling of production areas aid in determining of exact requirement and time of needed attention for pests control and field roguing of the stock. The samplings are made on plans or designs which permit significant difference determinations. Following production is the harvesting and handling of the unmilled seed for temporary storage, which calls for constant care throughout to maintain product quality. This is achieved primarily by trained supervisors.

The milling of the seed crop is a long procedure requiring annually five to six months to complete the job. Since these crops were grown by many farmers on firm price contract basis and in widely distant agricultural regions, some common type of buyer settlement based on a mutually acceptable seed quality standard is employed. The purchase settlement is based on the net clean seed obtained from the milling operation. It is this operation which requires quality control of the seed product in order that the quality output is consistent from one day to the next, from one field lot to another, and for each of the several milling machines in each seedhouse and from one area seedhouse to one in another location.

Milling samples are drawn hourly from each operating machine, and equipment is kept in adjustment to maintain the required quality level. The finished seed is warehoused in sequence of its milling and final quality designations are made after vitality tests have been completed. Other qualitative characteristics of the seed and its container (bag) are made a part of the final seed quality report, lot by lot.

Following shipping of the seed stocks to the far flung vegetable crop production areas for planting, all lots are again sampled, tested and the results compared to the original evaluation when each was adjudged as being of acceptable quality level. When the results of these two tests are found comparable, the seed is then approved as seed planting stock for the production of the Company's canning and freezing crops.

All these SQC applications in the seed development program are a necessary ingredient in the total program, as this makes possible the most uniform seed quality and thus a relatively uniform raw product.

As the pea seed is planted in preparation for the green pea harvest, a close analysis of weather conditions aids greatly in continued control of the process. The amount and rate of heat and moisture have particular effect on the uniformity of development throughout the period of production and harvest. Control charts set up for heat measurement (usually termed "heat-units") and moisture measurements will readily tell whether these important weather conditions are normal, high or low. This information in turn tells us whether to seed at the normal rate, faster than normal, or slower than normal; and as the crop matures, this procedure indicates how rapidly the crop is approaching the harvest date. Since the critical criterion for harvest of green peas is the maturity factor, an additional measure of harvest date is usually scheduled for this maturity consideration. Routine "pre-sampling" from the field a few days in advance of expected harvest with proper analysis of the progressive maturity grades can provide a very positive indication for optimum harvest date.

As we progress on to the actual harvest, the critical quality consideration remains as the maturity or texture characteristic. This maturity factor is generally evaluated by measurement on some type of texture meter such as tenderometer or shearpress. The maturity data collected at this point is used for three distinct purposes. The first purpose being a basis for establishing purchase price, the second purpose as basis for scheduling each lot efficiently into the plant operation, and the third purpose as basis for control of the continuing harvest from the particular field in question. Routine control charts will often be helpful to aid in accurately estimating the maturity level at time of harvest, and in all cases it is extremely important to know the normal within-sample and between-sample variance. At this point of harvest it may often be desirable to analyze other quality factors such as extraneous vegetable material or broken peas. The greatest value in this analysis is likely for efficient scheduling of raw product lots into plant production. It should be noted here that the analysis technique applied for the raw product stage of production may often be somewhat of a time consuming venture at the outset. However, it will also be found that mother nature provides for a fairly stabilized pattern of variability, such that the variance estimated from a single well designed experiment should prove to be adequate basis for decision for an indefinite period. Simple routine applications may then be found practical.

The basic estimates for evaluation of raw product data for peas, as applied today at Stokely-Van Camp, are based on calculations made several years ago.

It may be in order at this point to emphasize more clearly the consideration for efficient scheduling of raw peas into production. Perhaps the most important direct economic potential for in-plant quality control activity falls in this field of endeavor. Certainly, any error in estimating the raw product grade for peas into production would likely suggest improper adjustment of mechanical or manual quality grading and sorting operations. Errors in the preliminary adjustment of these processes, even though they may be quite quickly detected and corrected, will produce appreciable quantities of finished production at the wrong quality level. Consideration must always be given to precision and accuracy of the maturity factor, in order to estimate the most efficient procedures for production of the particular raw product lot. However, most other quality factors are a critical consideration only intermittently.

If we have successfully obtained uniform raw product and have successfully scheduled this raw product into efficient plant production, we have made our plant production operation a relatively simple part of the entire process. Although, at this point of plant production, there is potential for the greatest number of SQC applications in the entire process, most applications here are relatively routine and simple control charts applied to all the important quality factors.

The immediate need for efficient SQC applications in the plant production operation may be assigned to the fact that at this point the final concentration of effort is made to culminate the entire lengthy production phase. Failure, at this point, to fully utilize the raw product available and thus efficiently obtain the potential recovery and grade yield, shall certainly contribute significantly to high cost for end product. In addition, there is the obvious fact that nothing beyond this phase of the process can further improve the product quality. The many quality factors of importance to the finished product (i.e. maturity, broken peas, blemished peas, off-color peas, extraneous material, blanch, package fill weight, labeling, freezing and sanitation) all emphasize this potential for SQC applications in the plant. Efficient control of all these factors is extremely important at the time raw product is placed into production and at each subsequent point until final movement of the peas into storage. Since this plant production stage of the process presents the most obvious need for effective control, initial widespread applications of SQC in frozen pea production at Stokely's were made at this point. The control chart application at the point of quality grading of peas for the maturity factor is one of the most effective demonstrations I have seen for substantiating the value of this technique. Although the agronomist is correctly credited with the early development of statistical philosophy and our Seed Division followed that pattern for many years, it was successful widespread application of SQC techniques in the plant that emphasized the overall potential and ultimate integrated program.

As throughout, the maturity factor is most important in the plant production phase of the frozen pea process. This maturity level is adjusted by quality grading (separation of different maturity levels) the peas in salt brine solution. Routine control chart analysis of quality data describing this quality grading operation permits most efficient basis for control of this factor. Simple routine techniques may also be applied to evaluate the efficiency of this quality grading operation at all times. Similarly the mechanical cleaning operations are controlled with aid of control charts in order to efficiently remove extraneous material and broken peas. Although little can be done to quality grade or sort discolored or damaged peas, the defective control chart here provides a means for effective segregation of different quality levels. In addition, defective control charts quickly suggest to us whether the production flow should be sorted at normal production rates, at reduced production rate or adjusted quality grading conditions, or run intentionally as second grade product. The value of control charts as an aid in controlling the filling and labeling operations is quite widely recognized. Control chart application at the

filling and packaging operation was the initial in-plant SQC application at Stokely-Van Camp, as it has been for many other food processors.

Aside from the routine control chart applications for product quality guides, the special operations of blanching, freezing and sanitation or housekeeping (bacteriological data) all lend themselves to valuable routine control chart applications. In fact, failure to control these special functions accurately may well negate all previous effort by greatly hastening the rate of product deterioration or other costly failures. There have necessarily been many lengthy studies and utilization of numerous statistical methods of analysis as preliminary basis for these many simple, routine applications. And, of course, these more complex analyses are continued as the need dictates.

It should be apparent that efficient control of the production and manufacturing operations for the frozen pea process requires a long range program of planning and production scheduling. Efficient scheduling of production should consider sound statistical analysis of past history for variables such as plant production rate, raw product recovery, pack out by grade, pack starting date, pack ending date, market distribution pattern for finished goods, and most important, accurate sales forecasts. The particular planning and scheduling problem may involve a single plant operation or multiplant operation, a limited sales area or a vast complex of sales area, a one crop season or multi-crop season. The frozen pea process at Stokely-Van Camp would involve considerations for multi-plant operations, complex sales areas and single crop season.

In forecasting sales for our frozen pea pack, we must consider several main points. These considerations involve knowledge of sales history by item, by customer, by marketing area, by labels distributed nationally, by labels distributed regionally, by first grade labels, by second grade labels, by retail size containers and by institutional or bulk size containers. The various statistical techniques commonly used for sales forecasting are utilized throughout the sales analysis and estimating process in an effort to obtain a final best estimate which is based on the facts of the case, and thus non-biased. Initial consideration should be given to calculation of the regression curve to establish the trend and current level of sales. This regression analysis also provides a sound basis for quickly estimating whether or not a significant shift of sales pattern has taken place. Experience suggests that monthly sales data for a twenty-four month period generally is proper basis for this estimate. The monthly data is adjusted for seasonal variation and trend.

After reviewing the facts of the case as expressed by the statistical analysis, sales management must confer on both the tangible and intangible aspects of this data. General business trends and fluctuation are considered as they may influence our objectives; all aspects of pricing, warehousing and traffic flow are considered; and knowledge of customer prospects in adding (or dropping) new sales outlets are taken into the final estimate.

Basic to all of the entire process is the Quality Control Department, whose primary function is to assist management in defining, establishing and maintaining standards of quality. The quality standards, thus established, must be a primary consideration for planning the proper varieties, production locations and methods of production. An important aid to developing competitive standards is utilizing regular test panel analysis of market samples of various competitive brands. Test panel analysis techniques are also valuable aids in choosing between varieties or processing differences. Similar techniques are utilized for estimating consumer reaction to existing quality standards, to new products or contemplated changes in existing products.

Numerous SQC applications are utilized on an irregular basis for purposes of determining degree of correlation between different factors or test methods, coordinating functions between plants, work sampling studies, testing of supplies, checking clerical accuracy (or merely sampling clerical data in lieu of 100% sampling), special research projects, acceptance sampling and many other functions.

Although this particular frozen pea process has suggested that a balanced SQC program be appropriately a near completely integrated SQC program, other food production, manufacturing and marketing programs may well require emphasis on particular phases of the process in order to arrive at proper balance and thus the optimum economic effect.

If we consider each phase of the entire process to be one component of a many component function, we can estimate that the overall efficiency of the entire process is equal to the product of the efficiency of each phase. Our conclusion then would be that we must strive for the most efficient operation at each component phase in order to achieve a relatively high degree of efficiency for the total process. Since no manager wants his department to be recognized as inefficient, it should not be an impossible task to emphasize this point. Scientific SQC techniques properly applied at any phase of the process should produce the most efficient operation. Thus, it would seem apparent that balanced SQC should be a primary goal for any food processing operation that intends to work toward an overall high degree of efficiency.

A review of the frozen pea operation at Stokely-Van Camp would reveal that many of the standard statistical techniques have been utilized effectively in all phases of the process. This process approaches completely integrated SQC, although not always in a strict theoretical sense. In fact, it is rare that theoretical philosophy can be applied directly to the many routine SQC applications in food processing operations such as our frozen pea example. Many applications discussed here involve practical modification of theoretical considerations in order to provide economically feasible application. Many potential SQC applications may never justify the cost and are thus eliminated from consideration. Throughout all phases of our operations there are continuing efforts to at least informally establish appropriate "bench marks" or standards of quality, and always there is an appreciation and knowledge of "Sigma". Those of you who are familiar with the successful application of Linear Programing for the solution of the overall planning aspect, will readily recognize the direction for future potential development of these scientific techniques in our frozen pea example. This most direct and efficient application that could relate production, scheduling, manufacturing, sales, warehousing and shipping to a single general plan would, of course, greatly increase the efficiency of each phase of the entire process. In fact, this application would likely be the most effective application of all. Our particular process is progressing constantly in the ultimate direction of this linear programing operation. This phase in the development of our integrated program is proceeding as the various individuals involved become more familiar and appreciative of the scientific methods and philosophy. As in even the most elementary SQC applications, this "educational process" has always been necessary.

I would assign, as the primary consideration, an appreciation of the very fundamental concepts of statistics. To acquaint all management personnel with these basic ideas should be the initial aim for quality control. Integrated SQC will then be the inevitable accomplishment.

SOME STATISTICAL CONCEPTS APPLIED TO THE NAVY CALIBRATION PROGRAM

Stanley Crandon
Field Operations Div., Metrology Department
Bureau of Naval Weapons Representative
Pomona, California

In this paper, I intend to describe some of the statistical concepts applied to the Navy Calibration Program. In addition, you will see how the application of one of these concepts has directly resulted in savings exceeding one million dollars per year. Before discussing these statistical methods, I am sure that a brief description of the program with some of its most significant features is in order.

About two years ago, the Navy's Calibration Program was actually a combination of individual programs administered by each of three technical bureaus: the Bureau of Ships, the Bureau of Ordnance, and the Bureau of Aeronautics. Each of the three programs was composed of a network of calibration laboratories which were known, somewhat interchangeably, as primary, reference, secondary, local, or tender standards laboratories. While this problem of nonstandardization in a standards program has not yet been completely resolved, other events have occurred which have considerably diminished the confusion in terminology.

The Bureau of Ships and the Bureau of Ordnance entered into an agreement which provided the following as some of its most important advantages: interbureau cross-servicing, joint support of a special technical research, development, and consultant group located at the Bureau of Naval Weapons Representative, Pomona, California, and an integrated calibration procedure preparation effort. This became known as the BuOrd-BuShips Calibration Program. While this was now a joint program, each of the two bureaus retained certain distinctive characteristics. For example, the Bureau of Ordnance provides calibration services for BuOrd prime contractors at the secondary standards laboratory level and for all Navy prime contractors at the primary standards laboratory level; the Bureau of Ships program does not offer any calibration services to contractors. Another basic difference lies in the measurement coverage; the BuShips program applies only to electrical-electronic areas, whereas the BuOrd program encompasses all measurement areas. The BuOrd-BuShips Calibration Program could be described as a single program under dual administration.

Just a few months ago, the Bureau of Ordnance and the Bureau of Aeronautics were combined to form the Bureau of Naval Weapons. The most immediate effect of this merger upon the BuOrd-BuShips program was a change in name to the BuWeps-BuShips Calibration Program. At this time, arrangements with respect to cross-servicing, laboratory echelons, technical guidance, etc. have not been finally completed by the Bureau of Naval Weapons. However, the anticipated laboratory echelon configuration is presented in Figure 1. In Figure 1, the laboratory nomenclature may be reconciled by comparing the number of times the laboratory is removed from the National Bureau of Standards. For example, a secondary standards laboratory and a reference standards laboratory are equivalent in common measurement areas as they are both twice removed from NBS.

NAVY CALIBRATION PROGRAM STRUCTURE

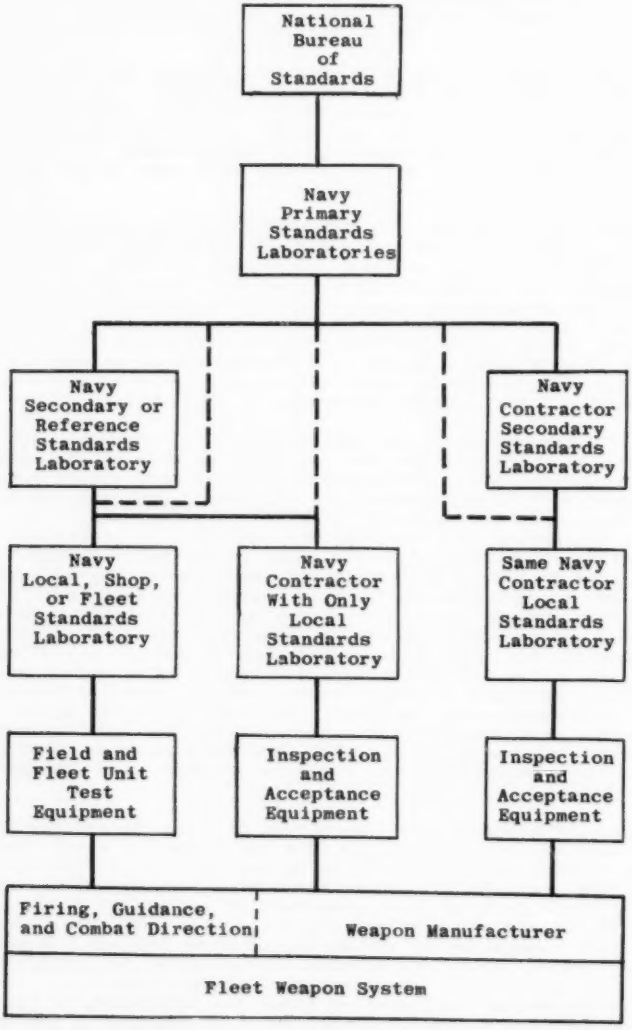

National Bureau of Standards
Navy Primary Standards Laboratories
Navy Secondary or Reference Standards Laboratory
Navy Contractor Secondary Standards Laboratory
Navy Local, Shop, or Fleet Standards Laboratory
Navy Contractor With Only Local Standards Laboratory
Same Navy Contractor Local Standards Laboratory
Field and Fleet Unit Test Equipment
Inspection and Acceptance Equipment
Inspection and Acceptance Equipment
Firing, Guidance, and Combat Direction
Weapon Manufacturer
Fleet Weapon System
Normal Calibration Flow
Alternate Flow Dictated by Accuracy Requirements

Figure 1. Navy Calibration Program Structure

Still another segment joined the Navy Calibration Program: the Fleet Ballistic Missile Weapon System (Polaris) Calibration Program. This program is generally in effect at FBMWS contractors' plants although some Navy field establishments are included; these are the Naval Weapons Annex at Charleston, South Carolina, where the Polaris missiles are assembled, NOTS China Lake, where the Polaris motors are tested, and the Atlantic Missile Range, Cape Canaveral, where the Polaris missiles are tested and fired. Some of the Polaris contractors maintain their own primary standards laboratory but most employ the services of the Navy primary or secondary standards laboratories.

Mention was made earlier of a special technical research, development, and consultant group located at the Bureau of Naval Weapons Representative, Pomona, California. This group, which is now known as the Metrology Department, has been in existence for over five years and performs the following tasks in the Calibration Program:

1. Recommends standards and associated measurement equipment to be used by standards laboratories, evaluating these to assure high quality selections commensurate with technical requirements.

2. Monitors measurement requirements and new developments in standards and associated measurement equipment and, where technical requirements dictate, develops and evaluates special calibration standards and fixtures.

3. Prepares and evaluates calibration, certification, and cross-check procedures and techniques at all laboratory levels.

4. Designates calibration, certification, and cross-check intervals for standards and associated measuring equipment.

5. Devises and establishes a quality assurance system for the program to include monitoring of all calibrations performed to assure compliance with calibration periods, determination of calibration quality and effectiveness of operating procedures, and investigation of operation and quality problems.

6. Provides engineering assistance in establishing new laboratories in the program, determining capabilities of prospective participants to accommodate measurement requirements and recommending additional systems as required. Designs laboratory facilities to solve equipment installation or similar problems peculiar to particular laboratories.

7. Designs necessary packaging and shipping systems to assure safe transit environment during shipment of standards.

8. Devises and implements training programs for standards laboratory personnel.

The Standards Laboratory Information Manual (SLIM) which documents most of the underlying concepts of the Navy Calibration Program was prepared by the Metrology Department, BuWepsRep, Pomona.

With this brief picture of the Navy Calibration Program, I would like to discuss a few of the statistical concepts utilized by the Metrology Department in performance of some of the above-mentioned tasks.

Some time ago, a BuOrd secondary standards laboratory DC voltage system (1.6 to 1500 volts) was established, utilizing a

potentiometer, standard cell, volt box, and associated devices. The accuracy specified for this system, ±0.045%, was considered adequate for almost all of the electrical DC meters likely to be encountered. Accordingly, calibration procedures utilizing this ±0.045% system were written. Notable exclusions from coverage by these procedures were such instruments as ±0.1% laboratory standards. Although there were apparently no calibration systems sufficiently accurate at the secondary level to perform satisfactory calibrations of ±0.1% instruments, these instruments could not be ignored indefinitely. The obvious action taken was to re-examine the existing system to determine if its accurary could be improved. In doing this, it was found that the ±0.045% accuracy was derived in the classical manner of adding the tolerances of each of the critical components of the system: ±0.015% for the potentiometer, ±0.01% for the standard cell, and ±0.02% for the volt box. The statistical approach, making the assumption of equal confidence coefficients for the three critical components and by means of partial differential equations, resulted in an overall accuracy of ±0.027%. This immediately opened the way to calibrate the rapidly increasing group of ±0.1% accuracy instruments.

The same approach was then applied to a DC current system (0.00001 to 1.0 amperes) and has reduced the classical estimate of accuracy of ±0.03% to ±0.018%. The Navy is presently re-evaluating all critical measurement system accuracy statements in light of this belated development.

Another application of statistical concepts is found in the design and analysis of measurement problems. Here, audit samples are sent to standards laboratories for measurement. Examples of audit samples prepared for primary level laboratories are:

1. Standard resistor, Thomas type
2. Standard capacitor
3. Standard inductor
4. Standard voltage source
5. Directional coupler
6. Precision attenuator
7. Coaxial attenuator
8. ASTM thermometer
9. Mass standard
10. Proving ring

Audit samples prepared for secondary or reference level laboratories may include, in addition to those above, AC-DC meter or decade type instruments, strain gages, aneroid barometers, etc. Although audit samples are usually selected for their high stability and resistance to both testing and transportation effects, an attempt is made to balance out possible effects by preparing packages in duplicate or triplicate and routing each package in a different order. This approach not only solves two problems but also provides additional measurements needed to reduce sampling error.

The statistical analysis of measurement problems is designed to test the null hypothesis that there are no significant differences in the measurements made by each laboratory. If this hypothesis is rejected, the analysis of variance will permit additional comparisons to be made between significant groupings of laboratories; e.g., former BuOrd primary laboratories versus former BuAer primary laboratories, secondary and reference laboratories serviced by the former BuOrd Western Primary Standards Laboratory versus those serviced by the former BuOrd Eastern Primary Standards Laboratory, etc. In the very unlikely event

that it cannot be resolved which laboratory's measurements are correct, referral is made to the National Bureau of Standards.

It was mentioned earlier that one of the tasks of the Metrology Department at the Bureau of Naval Weapons Representative, Pomona was to monitor all calibrations performed by particular activities. At the time of this writing, calibration data are being provided by 24 Navy activities and 25 Navy prime contractors (23 of which are Fleet Ballistic Missile contractors). It is anticipated that by the end of the year, 60 Navy activities will be submitting calibration data.

Most of the calibration data being provided are on a standardized form and consist of both attributes and variables type data; front and back portions of the standardized form are shown in Figures 2 and 3. Because of the magnitude of the data and the immensity of the task of reducing the variables data to common denominators which will lend themselves to EAM techniques of data reduction, the statistical analysis commences with the attributes data and advances to the variables data only if warranted by the results.

Referring to Figure 2, the information appearing in Sections A, B, and C is punched into EAM cards. After a preliminary listing of the cards, sorted according to calibration laboratory, submitting activity, manufacturer, and model number, in that order, the cards are resorted by manufacturer and model number into significant groups of interest. These may be BuOrd secondary laboratories, BuOrd primary laboratories, BuShips reference laboratories, Polaris contractors, etc. A second pass of the cards through a tabulating machine and electronic calculator will provide, for each group of interest and each member of each group, the following information:

1. The number of times a particular instrument (manufacturer and model number) was serviced during the time period under consideration

2. A breakdown of the above by reason for servicing, periodic calibration or functional failure

3. The condition of a particular instrument as measured visually: good, poor, etc.

4. The number of times a particular instrument was submitted in an out-of-tolerance condition

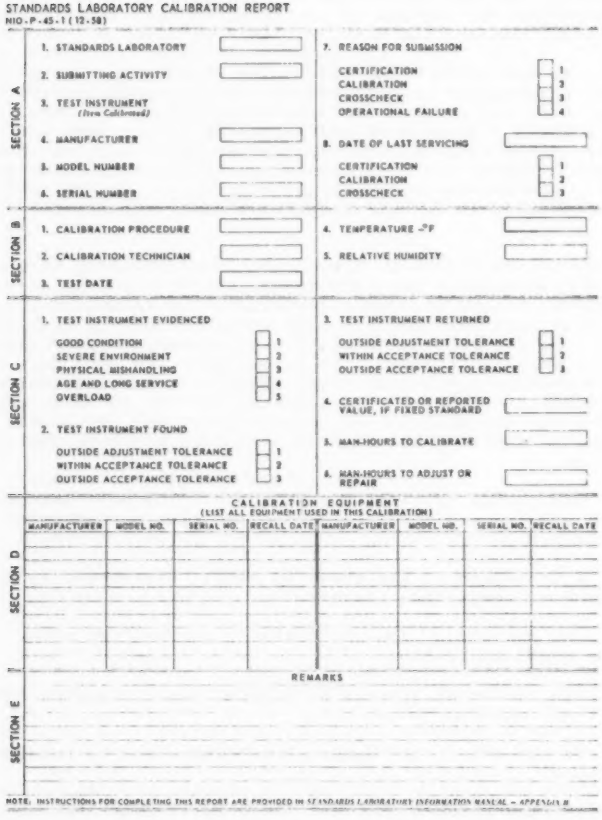
5. Of the above, the number which could not be returned in-tolerance

6. Average calibration time for a particular instrument

7. Average repair time for a particular instrument

Chi-square analysis is applied to these data to determine if there are significant differences between the various laboratories with respect to the attributes listed above. Reference is also made to the variables data on the calibration report forms, if considered necessary.

Perhaps the most significant and far-reaching result of these analyses has been the changing of instrument calibration recall periods. Our estimates suggest that since the original listing of instrument calibration periods in the Standards Laboratory Information Manual, 760 different instruments have had their calibration



STANDARDS LABORATORY CALIBRATION REPORT
NIO-P-45-1 (12-58)

SECTION A

1. STANDARDS LABORATORY
2. SUBMITTING ACTIVITY
3. TEST INSTRUMENT (Item Calibrated)
4. MANUFACTURER
5. MODEL NUMBER
6. SERIAL NUMBER
7. REASON FOR SUBMISSION
   CERTIFICATION 1
   CALIBRATION 2
   CROSSCHECK 3
   OPERATIONAL FAILURE 4
8. DATE OF LAST SERVICING
   CERTIFICATION 1
   CALIBRATION 2
   CROSSCHECK 3

SECTION B

1. CALIBRATION PROCEDURE
2. CALIBRATION TECHNICIAN
3. TEST DATE
4. TEMPERATURE °F
5. RELATIVE HUMIDITY

SECTION C

1. TEST INSTRUMENT EVIDENCED
   GOOD CONDITION 1
   SEVERE ENVIRONMENT 2
   PHYSICAL MISHANDLING 3
   AGE AND LONG SERVICE 4
   OVERLOAD 5
2. TEST INSTRUMENT FOUND
   OUTSIDE ADJUSTMENT TOLERANCE 1
   WITHIN ACCEPTANCE TOLERANCE 2
   OUTSIDE ACCEPTANCE TOLERANCE 3
3. TEST INSTRUMENT RETURNED
   OUTSIDE ADJUSTMENT TOLERANCE 1
   WITHIN ACCEPTANCE TOLERANCE 2
   OUTSIDE ACCEPTANCE TOLERANCE 3
4. CERTIFICATED OR REPORTED VALUE, IF FIXED STANDARD
5. MAN-HOURS TO CALIBRATE
6. MAN-HOURS TO ADJUST OR REPAIR

SECTION D

CALIBRATION EQUIPMENT
(LIST ALL EQUIPMENT USED IN THIS CALIBRATION)

MANUFACTURER	MODEL NO.	SERIAL NO.	RECALL DATE	MANUFACTURER	MODEL NO.	SERIAL NO.	RECALL DATE

SECTION E

REMARKS

NOTE: INSTRUCTIONS FOR COMPLETING THIS REPORT ARE PROVIDED IN *STANDARDS LABORATORY INFORMATION MANUAL — APPENDIX B*

Figure 2. Standards Laboratory Calibration Report (Front View)

SECTION F — CALIBRATION DATA

PROCEDURE STEP NO. (1)	FUNCTION TESTED (2)	READING OR VALUE		CORRECTED VALUE OF COLUMN ___ (5)	ALGEBRAIC ERROR (6)	CALIBRATION TOLERANCES		READING AFTER ADJUSTMENT (9)
		NOMINAL (3)	MEASURED (4)			ACCEPTANCE (7)	ADJUSTMENT (8)	

NOTE: INSTRUCTIONS FOR COMPLETING THIS REPORT ARE PROVIDED IN *STANDARDS LABORATORY INFORMATION MANUAL – APPENDIX B*

Figure 3. Standards Laboratory Calibration Report (Back View)

interval extended and 20 different instruments have had the calibration interval reduced. Based on an average reduction in calibration frequency from six to four times per year and an average calibration time of 2-1/2 hours, the annual dollar savings attributed to the above analysis exceeds one million dollars.

"HIGHER QUALITY" THROUGH "PEOPLE IN QUALITY - QUALITY IN PEOPLE"

Eugene V. Grumman, Chief Engineer
The Bullard Company, Bridgeport, Connecticut

Why the Need for Higher Quality?

Why do customer demands today place a very high emphasis on quality in nearly every purchase they make? Why does a constant search go on to seek out the better, more economic buy regardless of what field you may be in?

It is obvious that this is as it should be for everybody wants and needs the best for his class of product if he expects to meet competition.

We are now witnessing a second industrial revolution brought about by customer demands resulting in increased mechanization -- the use of man-made and man-conceived devices now applicable in one form or another to nearly every industrial situation. Such devices are necessary; more important, they are mandatory if we are to retain the lead of price, quality, productivity and delivery over competition regardless of its source.

Despite our attempts to equip ourselves with such devices, we will still be reliant on man himself to put them to maximum utilization in order to realthe maximum gains therefrom.

How can this be done?

I am not sure that there is any pat answer to this question. But let's attempt to find an answer to this question; let's take a look at -- where we are and what we can do about it, with particular emphasis on product quality.

The Basic Needs for Every Business

For every company that goes into business there are six basic requirements. It must have (1)

1. Management - to direct the organization
2. Men - to control and operate
3. Methods - profitable means of conversion
4. Machines - which convert
5. Materials - to be converted into products or services
6. Money - financial where-with-all to do business

Any one without the others makes the operation of a business an impossibility. A never to be forgotten point: Your competition has access to all of the things your company has -- except one thing - - - - - "Your People". This requires some serious thought for if a competitor succeeds over us, there can be but one basic reason - - - - - "People". Each of the six "M's" has a quality of its own -- but the success or failure of your company over competition is dependent on the quality of the capability and performance of your people.

What to do to Stay in Business

If your company wants to stay in business, it must earn a profit or it will soon cease to exist. With a given product line, there are six ways to earn profit: (2)

1. Raise prices
2. More sales
3. Cheapen the product
4. Redesign the product
5. Increase the workers' productivity
   (Management, Sales, Engineering,
   Manufacturing - service areas)

6. Eliminate waste
(manpower, materials, methods, money)

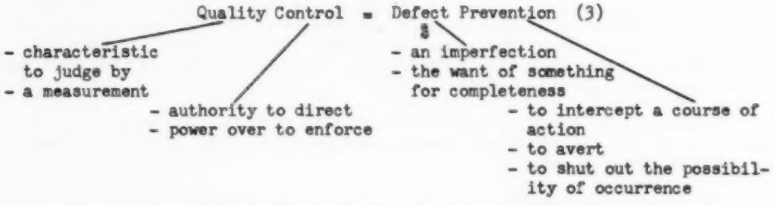
In short, if we want to earn a profit, we will do it only by making a quality product for less so that more people can afford to buy.

Regardless of how you look at it, for any company or system, the last two methods are the only means of increasing real profits.

As we examine Items 5 and 6 it becomes evident that the concept of quality control is right in the middle and becomes involved with #6 in particular. So let's take a look at quality control, what it can do and how some of these things were approached and accomplished.

Definition of Quality Control

The first and most important was to define quality control so that everyone knew what we were talking about.


Quality Control = Defect Prevention (3)
- characteristic to judge by - a measurement
- authority to direct - power over to enforce
- an imperfection - the want of something for completeness
- to intercept a course of action - to avert - to shut out the possibility of occurrence

This becomes quite interesting for, as we define the term, it points clearly to the fact that the authority to direct, to intercept imperfections in the measurable characteristics of the product rests in two places -- the machine operator and supervision. Quality must be built into the product right here.

Definition of the Three Levels of Quality

The second step was to clearly define the nature of the quality which must be controlled and to make clear the area responsible for it. (4)

1. Quality of Design (Engineering)
2. Quality of Conformance (Manufacturing)
3. Quality of Decision (Management)

What do we mean by quality of design?

We mean combining our past and present knowledge and experience with the known past and present experience of our customers -- plus, the known future requirements of both -- plus, creating new machines, new devices, mechanisms, controls and combinations of things -- which will result in greater productivity for our customer and irrespective of price makes our product a sound and an economical investment. This is the area of Management - Men - Materials, where Management is taken to mean the technique of getting things done through people.

What does quality of conformance mean to us?

We define the "quality of conformance" as the specific responsibility of the manufacturing division in that it shall produce the product to specifications as established by Engineering.

The part of quality control here is to provide the measuring devices, methods and techniques of auditing and measuring to assure management that manufacturing can accomplish this task. Further, to provide a systematic "feed back" analysis of audits, tolerances, machine and process capabilities that will permit engineering to do a realistic engineering job.

This area covers Men, Methods, Machines and Materials.

Now let's look at "quality of decision".

First off, we must recognize that when only one course of action is possible, then "no" decision is necessary. Only "if" there are alternatives can a decision be made. Alternatives sometimes lead us into complex situations -- put in another way it means this: Decisions are complex affairs. Complex problems have complexes.

Basically there are three ways by which we can make a decision:

1. Use a qualitative procedure.
   Use our authority, opinion, hunch, intuition, rumor, logic or be directed by past experience. (Snap decisions come into this category.) This method relies heavily on the integrity and authority of the decision maker.

2. Use a quantitative procedure.
   RECOGNIZE - that a decision has to be made.
   EVALUATE - collect all pertinent data relating to the problem. Do some research - but "get the facts".
   ANALYZE - formulate and quantify the collected data.
   DECIDE - the course of action from the quantitative data after considering all the alternatives.

3. Some combination of Steps 1 and 2.

Step 1 is comparable to blind flying without instrumentation. You have nothing to worry about unless you crash or run out of gas. Step 2 will cause us to make a decision based on measured fact, a quantitative method. Quality control to a large extent deals with this procedure. Step 3 is necessary, for industrial data that is free from bias is not always available; consequently, skilled decision makers invariably take this route.

A very important thing to remember here is the fact that people do not accept qualitative data as well as quantitative data; or stated in another way, people do not respond as readily to qualitative data as they do to quantitative data. (5)

Regardless of how a decision is made, we must "follow up" to test the validity of our decisions if we expect to make better decisions in the future.

IMPORTANT POINTS ABOUT DECISIONS (6)

1. Progress is built on decisions.
2. Decisions are forced by alternatives.
3. Alternatives provide facts.
4. Facts are acquired through research.
5. Research gives direction to change.
6. Changes involve thinking, planning and making decisions.
7. Decisions destroy information.
8. Decisions reduce doubt.
9. Resistance to change is a normal human reaction.
10. Decisions involve risk.
11. Risk implies that there is a possibility of failure.
12. Failure implies the need for statistical evaluation.
13. Decisions are complex affairs.
14. Complex problems have complexes.

If we are to progress, we must not only accept these truths, but we must live with them every day of our lives.

This, the area of decision, covers Management, Men and Money.

Align Your Facts, but First You Must Get Them

True facts are the most elusive of all things to pinpoint in the industrial situation. All too often what we think is the trouble is not bothering us at

all. "Pet peeves" coupled with biased interactions of problems and people, not only develop unnoticed, but more often than not lead us down long and arduous routes in our efforts to solve industrial problems.

Thanks to Pareto, there is a simple route to align the facts so that we can proceed with the solutions. Pareto's Law states, "that the greatest amount of wealth is concentrated in the fewest number of people". (7)

Translating Pareto's Law into the industrial situation it means: "The greatest dollar losses are concentrated in very few places". This method of analysis is simple and readily understood by management. Frequent use of this technique to determine where the greatest losses occur, why they occur, which ones to work on, who is doing it and other related reasons will pay dividends throughout any quality control program.

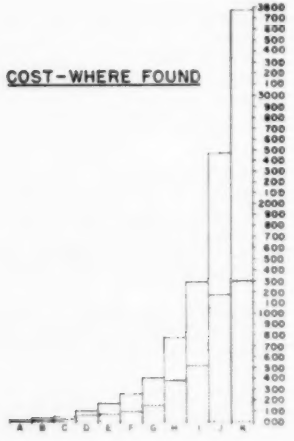


FIGURE 1

FIGURE 2

Management has long recognized that sizeable losses occurring in the industrial situation were avoidable and in its desire to do something about it has accepted the establishment of quality control as the true fact-finding media in the industrial situation.

Pre-Quality Control Set-Up in Industry

Let's take a brief look at the pre-quality control setup in industry that brought this about. (8) For the most part "so called" 100% inspection was frequently employed. This system at best is about 85% effective and direct examination showed that in many cases it rejected nearly as much good work as it did bad work. Such a system endeavored to expose the bad work -- a negative sort of approach. Further, and not by choice, through moral suasion or other pressures present in the industrial climate, it became responsible for product quality, yet had little or no authority to do anything about it. The caliber of people inducted into these areas during periods of rapid expansion left much to be desired.

Patrol inspection was employed and this in many cases proved haphazard and superficial. Failure to back up these people when a scrap producing machine was tagged, resulted in a degenerate inspection action through no fault of the inspector, but rather in a weakness in the system under which it operated.

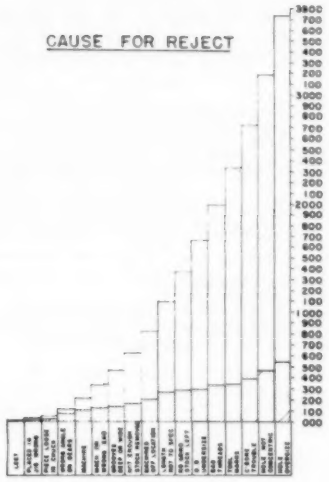


FIGURE 3

Systems where no formal inspection was in operation have been tried. Such a system relies on the integrity of the worker and/or supervisor. When the skill of the workers is high and the interactions between workers and supervisors is good, this type of primative inspection system can be difficult to beat.

In general, past inspection systems have not succeeded in raising product quality or in appreciably reducing avoidable losses. You cannot inspect quality into the product. Frequently, the examination of these systems showed the cost of inspection labor to be twice that of the loss due to defective products. (9) Further, the addition of inspection, in the long run, has little effect on productivity and rarely, if ever, eliminates customers' complaints to the extent desired.

FIGURE 4

Avoidable Losses Attributable to the Operator

Examination of records for many companies shows that losses in dollars for each worker range from $500 to $1000 per year. (10) This amounts to 25¢ to 50¢ for each hour he works. (Forty hours per week for 50 weeks per year.)

Examination of our own records revealed that we were below the midpoint of this figure. A 3-year historical record of defective work was available for study. Analysis of these records indicated that there was a lack of attention, a lack of thoroughness exhibited by the machine operator. Basically, the operator was not "conscientiously aware" of his responsibility to produce the product to specification.

Who is Really at Fault?

Under the former inspection system the operator frequently received instruc-

tions from the inspector and at the same time challenged him to find the substandard work. Operating under an incentive system, it would be to the operator's advantage if defective material passed inspection. By nature man is a gambler -- being a gambler, he is willing to take chances when there is no danger unto himself.

If the worker in the industrial situation can get away with making substandard material, we have to ask ourselves: "Is it the man alone that is at fault, or is it the system under which he works"? Fundamentally all people are honest and will remain that way unless we make it easy for them to be otherwise.

The 14% accounted for defective raw materials, engineering changes, improper and defective tools, wrong instructions and machine failures in that order.

Needed - A Closer Look at the Operator - Why is He like He is?

Since 85% of our troubles in terms of loss due to defects stems from the operator, it becomes evident that we must know something about why he is like he is. Our first step was one of educating ourselves to be cognizant of human needs. (11)

PSYCHOLOGICAL FACTORS

Basic Drive (Urge)	Want (Need)	Obstacle (Barrier)
Survival	Food, clothing, shelter, procreation	Unemployment
Security	Safety, freedom	Health, accident
Belonging	To be part of what is going on, identity	To be shut out, inability to attain social satisfaction
Self-esteem	Status, self-respect	Job classification as he sees it
Self-improvement	Work, recognition, a chance	What he does to make a living

We recognized that any change to be made in the industrial situation must satisfy the last three of the basic five drives and that such changes must not cause anxiety or frustration, but must allow for an improvement in the worker's attitude.

So let's define these things as they fit into the industrial situation.

Anxiety: Any uneasiness of mind over an impending or anticipated situation -- generally something in the future. (12)

Anxiety: is a vague fear
is enduring
results from conflicting needs
produces hostility

Frustration: A result of the obstacles real or imagined which prevent us from attaining a purpose or goal. Frustration occurs when the individual's motivated behavior finds a barrier in the way -- frustrated behavior is not directed intelligently. (13)

Attitude: All that a man reflects to us -- due to the significance that he attaches to the meanings and feelings aroused by a given situation. (14) Attitudes are set into motion by man's emotional overtones. Attitudes reflect a definite direction and can be seen as

constructive or destructive
helpful or harmful
positive or negative
compatible or hostile
and many others.

Another point we must consider, is the worker's emotional reaction and adjustment to the industrial environment -- the morale in his work area. This is important because low morale generally, but not always, means low productivity and frequently will be evident in the high dollar loss areas.

We should also look at each area to see how the workers are motivated. (15) All too often the industrial supervisor will define motivation as making the worker do what you want him to do. Let's not kid ourselves; our attempts to motivate others by the constant use of "demand directives", "command decisions", or the use of "threats" (all of which get quick action and are necessary under certain conditions) are not conducive to getting people to do things willingly or efficiently over the long run. Criticism and constant censoring have the same effects. You will find motivating people by threats, not only short in duration, but quite long in terms of quiescent retaliation. It is evident that those who talk about motivation most seem to know the least about it.

Motivation may be defined as "that within the individual which impels him to attain a goal or accomplish a task".

Dr. Glenn Gillman of Georgia Tech. expressed it beautifully as follows: "Motivation is a socio-psychological mechanism that moves us to satisfy a certain need in one way rather than another, in such a fashion that each step toward satisfaction is a meaningful step". (16)

It must be evident "to get that within the individual which will impel" is an important industrial task. The only way to do this is through a planned audio-visual approach.

After analysis of all possible points, we decided that there were three approaches we could use as a means of reducing defects, viz., statistical, practical and psychological.

Use Simple Statistical Procedures

We are a job shop. Lot sizes range from 1 to 25 pieces with occasional lots equalling 500 pieces. $\bar{X}$ & R Charts were tried and were not too satisfactory. It is difficult to get the industrial worker to pick up this concept for short run operation. Pre-control had not been developed when we started our quality control program. A simple fraction defective chart was devised using Mosteller and Tukey's binomial probability paper as a base. (17) This was quite successful since mathematical work was not required of the operator.

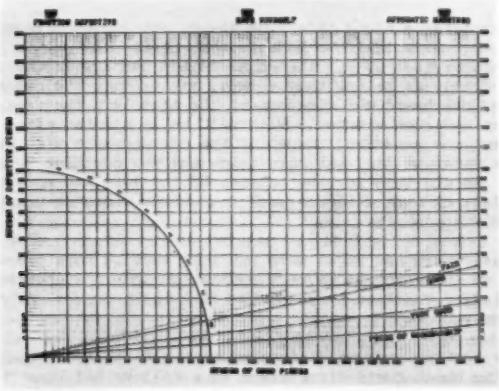


FIGURE 5

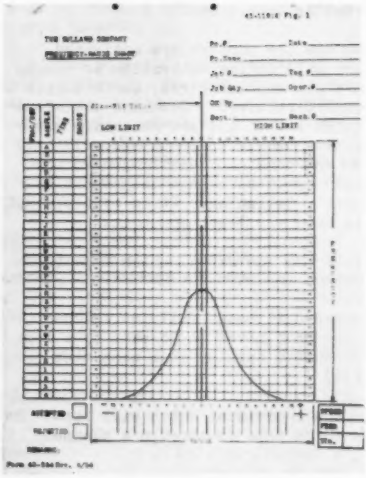


FIGURE 6

A simple chart which we call the Frequency Range Chart was developed. (18) Supervisors and machine operators could use this to make histograms of a process. Once done, quality control could convert this into statistical analysis to determine machine or process capabilities.

Pareto curve analysis was carried on by quality control to give supervision vital information on what caused the defects, who was doing it and where it was done.

We did not press too strongly for the use of statistical methods since it was quite evident that considerable training of both supervision and machine operators is necessary if these procedures are to be used effectively.

Be Practical - Get to the Root of the Problem - Arouse Curiosity

The practical approach was one of providing the most efficient measuring device right at the machine. In all cases variable measuring devices were selected. This provides the operator with a "compass" to control the direction of size for the parts which he is making.

We were concerned as to the best approach to get the operator to use these devices. To be successful the operator must be on our side. I am sure everyone, at some time or another, has seen Frank Buck's movie "Bring 'em Back Alive". You perhaps recall how he caught the monkeys by cutting a hole in a coconut, tieing the coconut to a stake and then placing a spoonful of rice in the coconut. The "curious monkeys" would stick their hands into the coconut to get a handful of rice, but could not get their hands out while holding onto the rice. The natives would then grab them and put them in a cage.

Our first experience along this line was to set variable gaging equipment and have the supervisor work it to be sure that it did the job to his satisfaction. With the supervisor we put the gages on the job and in the presence of the operator talked freely as to how they worked, making sure the operator witnessed the operation and overheard our discussion. We then left the area. The "monkey curiosity" in man was soon aroused. Within 10 to 15 minutes after we left, the operator was using the gages and comparing the measurements with his micrometer. About 30 minutes later he laid his mike aside and was checking everything by the variable gages. He did this for the balance of the lot, but more than this he did something else, he separated his work into three distinct groups -- acceptable, questionable and scrap.

In every case where new gages were put into service this procedure worked smoothly and in no instance did we have objections from the operator. We had as you can see satisfied the last three of the five basic drives and did not cause anxieties or frustration.

In our gear cutting area we discovered that all operators would walk to the inspection area to check their first piece on a Fellows Red Line Checker. Examination of the routine showed that the group of operators earned $4,000 of their combined yearly pay walking back and forth. A new red liner was purchased at a cost of $3,500 and placed in the machine area. Further examination revealed

this to be true in the gear hobbing area and a second red liner was purchased.

Close examination of equipment in the cutter grinding area revealed some antiquated measuring procedures. A 36" Optical Projector was purchased and used by the supervisors and skilled cutter grinders. The optical unit cost approximately $10,000 and replaced two inspectors; each earned better than $5,000 per year.

The optical unit caused us to spend additional sums on arbors, fixtures and index devices. This, because the equipment available was inadequate to permit us to get the type of grind and tolerances we required in our work. I urge you to look carefully in the tool and cutter grinding area for many quality problems start right here.

Observation in some of our milling operations revealed that excessive time was spent by the operator going to a surface plate to check work. Here is an excellent place to take the mountain to Mohammed -- put your surface plates on wheels, have the operator use it at the machine when necessary.

We had been troubled for some time with failure of taper ground bores in boring bars and spindles. Records revealed a loss of nearly $10,000 dollars a year due to these failures. Gaging by an attribute method there was always some doubt as to when a part was right and when it was not right. Attribute gaging methods were inconsistent even in the hands of experienced individuals. This is the qualitative system.

After disagreement with the gage manufacturers as to how to make a foolproof gage, we designed our own tapered air gage. We later purchased it from a commercial gage manufacturer at a cost of $5,000 and have not lost a single taper bore in seven years.

Another problem which bothered us was hardness testing. Many portable devices are currently on the market and pay dividends in our type of operation. I urge you to use this quantitative means right at the machine. More often than not, machining problems stem from micro structure rather than hardness.

Many other things similar in nature were done and new problems stick their heads up all the time. If we are alert, we will see them and take corrective action.

In a two-year period nearly $200,000 was spent to provide variable gaging equipment and services to take care of our manufacturing measuring requirements.

Need for Surveillance over Gages and Tools

With such a program we recognized that a reliable gage control area to maintain surveillance over this equipment was necessary. A temperature controlled area was provided, staffed and equipped with measuring machines for calibration. Further, a mobile service was inaugurated to inspect all gages, test bars, micrometers, etc., issued from the tool cribs. The latter was an eye opener when we started, for it found 27% of the measuring devices checked in the first three months were defective. Today this figure is less than .5%.

The gage control area is responsible to identify, catalog and inventory all measuring devices owned by the company. All variable gaging devices issued by gage control are signed for before release to fabrication areas.

Another important function was "Tool Cycle Inspection". It is pretty evident that the gages we give a man to measure his work don't mean too much unless the tools he uses are equally in good order. Cycle inspection of jigs and fixtures was another eye opener for 30% of all tools, jigs, fixtures inspected the first three months were defective. After the second or third go around these inspections can become random for the degree of defectiveness reduces to 1% or less and then is confined to reamers, drill bushings, registry points and the like.

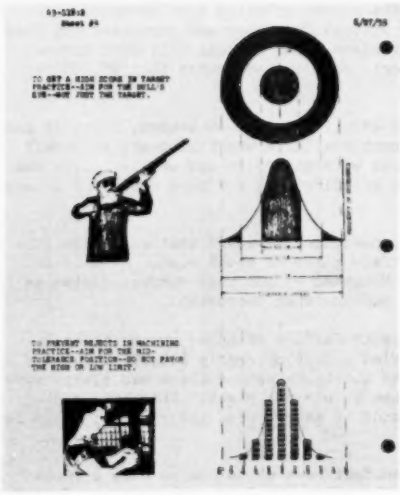


FIGURE 7

"IT IS NOT THE PURPOSE OF INSPECTION TO RELIEVE THE OPERATOR OR SUPERVISION OF THEIR RESPONSIBILITY FOR PRODUCING A QUALITY PRODUCT, BUT RATHER TO CHECK THEIR EFFECTIVENESS IN THE ACCOMPLISHMENT OF THE FOREGOING".

FIGURE 8

Use a Practical Psychological Approach - Aid Supervision

While all of the aforementioned things were in progress, the psychological part of the program did not remain still. One of the first things was to make use of the house organ to keep our people posted. This was done through short statements, pictures and easily understood charts. If you have something good, advertise -- tell your people about it. Communication is a very important part of the quality control responsibility.

A systematic procedure of returning records to supervision showing the "dollar loss" due to rejected or salvaged material, frequently -- almost always -- will spark a supervisor into action to eliminate such losses in the future. More than this, it will cause the supervisor to work more closely with his men to take planned preventive action.

It is quite important, I think, that quality control people learn more about industrial behavior, more about the psychological understanding of people in the industrial situation. I say this because supervision needs help on this score and there is no better place to get it than right on the floor. We sat in with supervision, when defects were discussed with the operator and gave every assistance possible. When you realize that thought concentration and pressure to meet production schedules is generally foremost in the supervisor's mind, you will understand why it is that you can render a much appreciated service in this area. I have as yet to see a supervisor resent this help. Remember motivation -- the audio-visual approach -- fits in right here.

A word of caution -- don't start out as the self-appointed expert. (19) Know yourself first. I urge you to read Dr. Reilly's book "Successful Human Relations"; then put it into action step by step.

The human engineers talk about part learning vs. whole learning. Part learning is a step by step process. Whole learning means to "shoot the works". Unfortunately, we do not have the ability to absorb nor retain when we "hear and see" too much at a given time. It has been proven by actual test that we rarely retain over 10% of what we read or hear at a given time. Planned positive steps produce the best results.

Will Your Efforts Pay Off?

Throughout this program a gradual reduction of inspection -- a nonproductive function -- took place. In some areas additional supervision was added. In our first year we went from 134 to 92 people. During this period the work

load remained at the same general level.

In all areas of industry where the workers do not work against measured time standards, we must be on the watch for those conditions which satisfy "Parkinson's Law -- work expands so as to fill the time available for its completion". Stated in another way it suggests this, people do not do all that they are capable of; they do what they think is expected of them. (20) I urge you to look carefully for with the passage of time we tend, if not alert, to overman unmeasured work areas. (21)

Further reductions were made in inspection by applying statistical auditing techniques to all materials going to stock. In all, losses due to sub-standard materials have reduced about 33% on the average and the inspection labor force reduced by about 45% since the institution of quality control in our plant.

Expand the Use of Machine Capability Studies - One of Your Most Important Tools

Assembly and erecting areas were examined and many changes made. In particular, we started in September of 1953 to run machine capability studies on all machines being built and demonstrated for the customer. We believe that the customer buys a machine tool on the basis of what it will produce for him at a profit. It is evident then, in buying our machine, that he expects it to produce parts to his specifications.

Our experiences in the manufacturing area and corrections made led us to believe that machine capability studies on new machines would be much better if run in our plant than if run in the customer's plant.

We had been guilty prior to the institution of this procedure of sending machines in the field which were not producing parts to specifications. Corrections made at the customer's plant are not only costly, but are aggravating to both the customer and manufacturer.

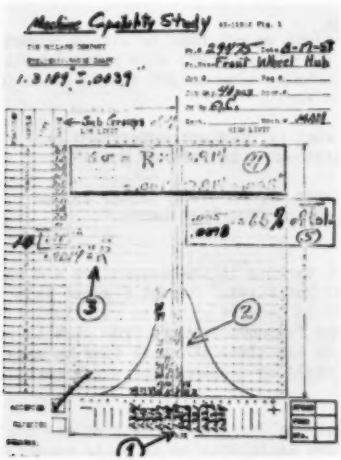


FIGURE 9

Our first studies were made using the $\bar{X}$ & R method. This is satisfactory when sufficient samples are available. For our purposes, we found a run of 30 to 50 pieces would give us all the information we required. It is desirable to keep the total number of samples small in order that tool wear, material variation, machine growth due to temperature changes, etc., are not major components of variation in the size of each part.

Other methods were used, viz., Sum and Range, Pre-Control, $\bar{R}d_2$ and Probability Paper developed by Art Bender. (22) In our first 160 studies we ran two methods simultaneously, $\bar{X}$ & R and $\bar{R}d_2$. The major difference was found in the fourth decimal place and this in less than 10% of the studies made. We continue to use $\bar{R}d_2$ when we have a minimum of 30 pieces for a run and use Pre-Control when we have 15 pieces or less.

In late 1955 we had the customer sending in teams of quality control people to make studies on the machines they were purchasing. We welcomed this approach. In most cases I feel it safe to say that most of

these groups knew a lot more about machine capability studies upon leaving than when they came to our plant. The exchange of ideas helped us to better understand the customer's problems.

Our objective when we started this service was to assure our management that the machine would perform to tolerance on the customer's part and to the customer's satisfaction.

Experience now, with something over 1,100 studies behind us, shows that our machines are capable of holding a natural tolerance of less than 75 to 80% of the customer's tolerance for the turning and boring work we do.

Today many of our customers accept the machine on the basis of the capability study made by us and mailed to them for their examination. We are quite proud of this for this was the ultimate goal we set out to accomplish when we established the service.

Many things show up during these studies. Some of these variables are listed here:

1. Improper alignment of slides and saddles
2. Improper gib adjustment
3. Securing screws bottoming
4. Faulty cutting speeds
5. Not cutting to size
6. Improper bearing adjustment in drive
7. Indexing errors
8. Chatter in surface finish
9. Machine vibration - fixtures out of balance
10. Improper feeds
11. Excessive chucking pressures
12. Out-of-round conditions

In all cases such conditions are either minimized or eliminated before passing the quality control study.

There is a Reason for Planned Positive Steps

At the start of our quality control program we met head on with a common quality control irritant known as "Acquiescent Resistance"; this was gradually changed to "Enthusiastic Participation". It takes time to get enthusiastic participants. Maybe the following will explain why:

> "The mind of man continuously rejects what it does not understand or that which is in conflict with the individual's beliefs or goals".
>
> "None of us are interested in changing our way of doing things unless we derive some inner satisfaction or benefit from the change."

Things to Come - A Capital Goods Problem

Today's production requirements, except perhaps in the automotive and related high volume industries, are in smaller lot sizes. Recent figures show 80% of the tonnage of all parts manufactured in this country are made in lot sizes of 20 pieces or less. (23) This means frequent changes from part to part. In many cases setup time alone may be four to six times the unit time to make the part. Such a condition does not lend itself to economic lot sizes much less to desirable profit margins.

The big problem facing us here is a real challenge and a test of the quality of our engineering and management decision capability -- to build numerical or other automatic controls which are simplified, of high quality, which create new markets for us, that permit fast changeover and higher productivity for our customer so that he can compete in his field -- if he does not profit, he cannot survive and in turn neither will capital goods survive.

What is the Answer to Our Questions?

We have covered a lot of ground up to this point and tried not to dwell on any particular phase of quality control. This was intentional for we were looking to find an answer to a question. At this point it must be obvious that the answer revolves around people. I think you will find that that answer lies in these areas -- in the physical and mental health and in the interactions of people in manufacturing, people in engineering, people in quality control, people in marketing, people in research, people in industrial psychology and people in the decision-making areas.

Think about it!

Where do you stand today?

Where will you be tomorrow?

REFERENCES *

(1) Churchman, Archoff & Arnoff - Introduction to Operation Research - Wiley

(2) American Management Association - Management Report #28 - Cost Reduction at Work - Page 7

(3) Author - definition used in Quality Control Training Programs

(4) Author - Bullard Quality Control Manual, Section 43-99:11

(5) Bridgeport University - Industrial Psychology #309

(6) Author - summation of lectures - Dr. Hugh Bowen, English Psychologist at Bridgeport University, and Don Fuller of Industrial Education Institute, Boston, Massachusetts

(7 & 10) Dr. J. M. Juran - Quality Control Handbook - McGraw Hill

(8) Author - Bullard Quality Control Manual, Section 43-99:10

(9) D. Shainin - lectures at University of Connecticut

(11 & 12) Smith - Psychology of Industrial Behavior - McGraw Hill

(13) Bernhardt - Practical Psychology - McGraw Hill

(14) Author - Human Relations Course for Industrial Supervision

(15) Giselli & Brown - Industrial Psychology - McGraw Hill

(16) ASQC - Automotive Section - 3/14/59 - University of Michigan, Ann Arbor

(17 & 18) Author - Bullard Quality Manual - Form 43-534-1

(19) Dr. William Reilly - Successful Human Relations - Harper

(20) Parker & Kleemier - Human Relations in Supervision - McGraw Hill

(21) Parkinson - Parkinson's Law - Houghton Mifflin

(22) Art Bender - Probability Paper - Delco Remy

(23) ASME - Numerical Control Seminar - New York City - November 1958

* Note: Reference material in this paper not necessarily taken verbatim from reference, but modified for better understanding and greater clarity.

AN APPROACH TO A RELIABILITY PREDICTION PROGRAM

N. E. Golovin
National Aeronautics and Space Administration

The purpose of the following remarks is to outline a point of view toward the reliability of complex systems which we have been developing in recent months. The approach is not altogether new; however, it may not have been adequately emphasized, partly because of extensive and somewhat fruitless searches for simple solutions to the problem. The general subject is of great current interest because of the skyrocketing importance of getting useful results in both military and civilian applications of space technology.

The immediate problems of reliability in systems of interest to us are naturally those of prediction. This condition reflects the fact that such systems are in their early R&D phases and there is practically no useful operating experience. Moreover, because of the high initial and later operating costs of the possible ensuing operational systems, the task of reliability prediction is an integral aspect of the forecasting of operational feasibility. In the case of satellite applications for purposes such as communications, for example, there is the additional problem that alternative approaches for meeting requirements are available and the forecast of operational feasibility for any such system must, therefore, be a relative one with respect to the integrated effects of cost and over-all system reliability in competitive alternatives.

Consider then a situation in which we have a detailed system design before us and let's address ourselves to the problem of developing a general technique for predicting its reliability. Our procedure will be to develop a theoretical approach to the problem interspersed with some comments and comparisons related to current approaches in handling systems of arbitrary complexity.

From some points of view, the crux of the problem in such an analysis is two-fold: First, the matter of how one defines "failure," and second, how one attempts to construct an expression for the over-all reliability of a system.

Conventionally one considers two types of failure, the so-called "catastrophic" and "degradation" kinds. The first is associated with the sudden, total failure of an object of interest. The "degradation" type corresponds to gradual deterioration of one or more of an object's characteristics parameters to the point in time where an essential function can no longer be fully performed. In a fully valid analysis, it is, of course, necessary to maintain a continuing distinction between these types of failure. For our purposes here, however, it is really not necessary to do this. Accordingly, in the subsequent remarks we will combine these two types into one. We will say that an object fails at the time that any of its relevant physical characteristics attain values outside a specified range. Our analysis will then try to show how this range must be determined for a general approach to be consistent and useful.

As to the manner of constructing an expression for the over-all reliability of a system, the usual procedure is to begin with failure studies of parts and to construct from such data, successively, estimates for the reliability of components, subsystems, and finally of the system as a whole. We will reverse this usual procedure and start with a definition of failure for the system, and then work back through subsystems and components to the data on parts failures. The basic reason for this reversal in approach is a somewhat theoretical one; namely, the fact that a part cannot logically be said to have failed unless the over-all system has done so, and, in fact, parts can fail without appreciable effect on system performance. This means that

the definition of part failure must be completely implied by the quantitative definition of what constitutes system failure. This point of view is that adopted in MIL-STD-441 for the Reliability of Military Electronics Equipment(1), which suggests that the required performance of system details should be obtained by working back from functional requirements for the over-all system.

Now the system design must specify how its outputs must be distributed within certain specified ranges of values if the system's objectives are to be met. The failure of a system to meet design objectives is, therefore, generally subject to unambiguous, quantitative definition. Furthermore, considering the assembly of distinct subsystems which interact to insure the desired output characteristics of the system, we can also take as given a set of mathematical relationships which allow calculation of over-all system outputs from the characteristic outputs of the constituent subsystems. For a design to be at all realizable, such mathematical relationships are either deducible from applicable physical theory or can be empirically established from related prior experience with similar equipments. This is necessarily the case if subsystem, and lower order, nominal performance specification are to have a rational scientific foundation. As a matter of fact, if such is not clearly the case, it can probably be cogently argued that the state of the applicable theoretical and practical arts does not justify a major system development program.

The logically obvious next step is, of course, the employment of such a theoretical or empirical, mathematical representation for constructing, on a computer of adequate capacity, a system simulation program in terms of the output characteristics of all of the system's individual subsystems. Computer-based system simulation, further, will then allow the systematic study of the effects on over-all system outputs of arbitrary variations in the structure of subsystem output characteristics. The results of this type of investigation, ideally, will then be the unambiguous specification of quantitative ranges for subsystem outputs, individually and/or in interdependent groups, which must be maintained if over-all system outputs are to be within the ranges defined by the tolerance requirements for system nonfailure.

To emphasize the point, the importance of proceeding from the tolerance limits on over-all system outputs to the mathematically implied maximum ranges of allowed variation in subsystem outputs is, principally, that one thereby obtains a valid quantitative definition of subsystem failures; and furthermore, that these definitions then permit equally valid specification of the probabilities of failure of particular subsystems, or of combinations of these into groups, if some are found not to be individually independent with respect to failure. Thus, the probability of failure of a particular independent subsystem is the likelihood that one or more of its outputs fall outside the tolerance limits established as acceptable by such a computer-based simulation; and, similarly, the probability of failure of statistically interdependent groups of subsystems is the likelihood that the structure of the groups' outputs to other individual subsystems, or groups, falls outside the ranges specified by the simulation study. This quantitatively justified grouping of subsystems into statistically independent entities with respect to probabilities of success or failure will be a particularly useful by-product of the simulation analysis; the analysis will define the probabilities of success or failure for subsystem groupings to which the conventional "product rule" can be validly applied.

The remainder of the argument is clear. In similar fashion, one next treats each subsystem as a mathematically structured assembly of component outputs, and then each component as a mathematically related group of parts outputs. Employing computer simulation, there are then developed quantitative criteria for failures of components and parts, as well as their valid groupings for purposes of combining probabilities of success or failure.

The essential product of these successive simulation studies is then, simply, two-fold: (1) We have estimates of the permitted range of parameter values for each part in the system, as required for over-all system output

acceptability; and (2) we have a quantitatively justified rather than an arbitrary basis for combining part failure probabilities to obtain, successively, such probabilities for components, subsystems, and the over-all system.

No mention has so far been made of the environmental stress problem and its relevance to this approach. Briefly, the indicated procedure will supply the quantitative tolerance requirements on all elements in the system which must be maintained, of course, under the anticipated environmental conditions of system operation.

The approach I have outlined has been developing fruitfully, particularly during the last two or three years, in a number of other organizations concerned with complex systems. The White Sands Missile Range (2), the Rand Corporation, Chance-Vought Aircraft, Convair-Pomona, and the Autonetics Division of North American Aviation have each, in some measure, adopted this approach. At the White Sands Missile Range, for example, the current effort consists of constructing a probabilistic simulation model for the NIKE-HERCULES System in order to develop the technique fully with the view to applying it subsequently to Army missile systems at the design stage. At Autonetics, the technique has been, in part, employed for the analysis and prediction of the reliability of the guidance subsystem for MINUTEMAN. While I am not familiar with any system to which the approach has been applied in its entirety, it seems that it has sufficiently solid logical merit to grow in importance, particularly in the case of systems in which the commitment to invest in a proposed design is greatly dependent on a realistic and objectively founded prediction of its reliability, and, furthermore, where the anticipated investments are so great that thorough, and therefore costly, reliability analyses have unquestionable managerial justification.

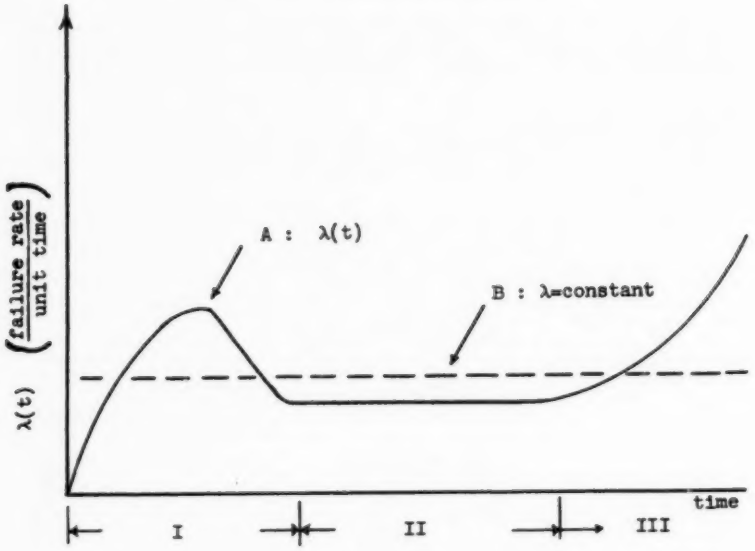
Let's complete our analysis by turning next to the question of how one develops a generally valid expression for the reliability of an individual part. For each part of the system our computer-simulation investigations have resulted in a quantitative specification of the ranges of acceptable variation in its parameters. Let's assume that we have as many parts as are required for an adequate sample, that we have a clearly defined environment in which the part will be required to maintain its characteristics, and that adequate facilities exist for carrying out a life test of the sample in such an environment. Incidentally, some of you may not know that of all the assumptions so far made in this discussion, these last are among the most unreasonable. Usually, at the design stage of a system, many parts do not yet exist, the environment in which they must operate is not clearly established, and testing facilities allowing study of their performance under a realistic reproduction of the anticipated environment are almost never available. We can test this sample until all of its members fail. The resulting data, combined with the usual assumptions justifying application of infinitesimal analysis, allows derivation of an expression, shown in Figure 1, for the reliability function, R(t), expressing the probability that the item survives the time t.

$$R(t) = R(o)e^{-\int_o^t \lambda(t)dt}$$

where $\lambda(t)$ represents the failure rate of the item per unit time.

The λ characteristic for the item may be arbitrary in character. In general, it is supposed to have the form of curve "A" shown in Figure 2.

Figure 1



Region I : 'Infant mortality" period

II : Mature operating life period

III : Rapid deterioration period

Figure 2

The straight line "B" is an average constant λ characteristic which, largely for the sake of simplicity in applications, is frequently assumed as applicable to most parts and components. This procedure is often justifiable as an initial approximation, or "first cut," at describing the corresponding reliability functions.

Substantial effort has gone into finding analytically tractable expressions for the part λ characteristic, or for its reciprocal defined as "the mean time to failure." As has been mentioned, the usual practices are to assume either that λ is constant, or that the part's mean time to failure is normally distributed about some average value with an appropriately chosen variance. There are many applications in which such simplified distributions are useful. However, it must be kept in mind that when many failure rates have to be added together to get a composite rate, the errors in such rates are also added. Accordingly, particularly in the case of complex systems, numerical methods allowing the use of actual rather than assumed part failure characteristics should be employed if at all possible. Additional reasons for care in this connection are suggested by the following:

1. In the first place, if a part is to be employed, without a prior "burn-in" or "warm-up" period, in a component in which it is duplicated a large number of times, the superposition of failure rates during the "infant mortality" periods may

result in a total failure rate, for an appreciable time, which is not acceptable for the component or subsystem. This effect is shown in Figure 3(a).

2. Secondly, if a number of different parts, with varying λ characteristics, are appropriately employed in a single component, it may be possible to easily design arrangements in critical circumstances that lead to a component λ characteristic which has desired form and a maximum allowable failure rate level. How this can be done is shown in Figure 3(b).

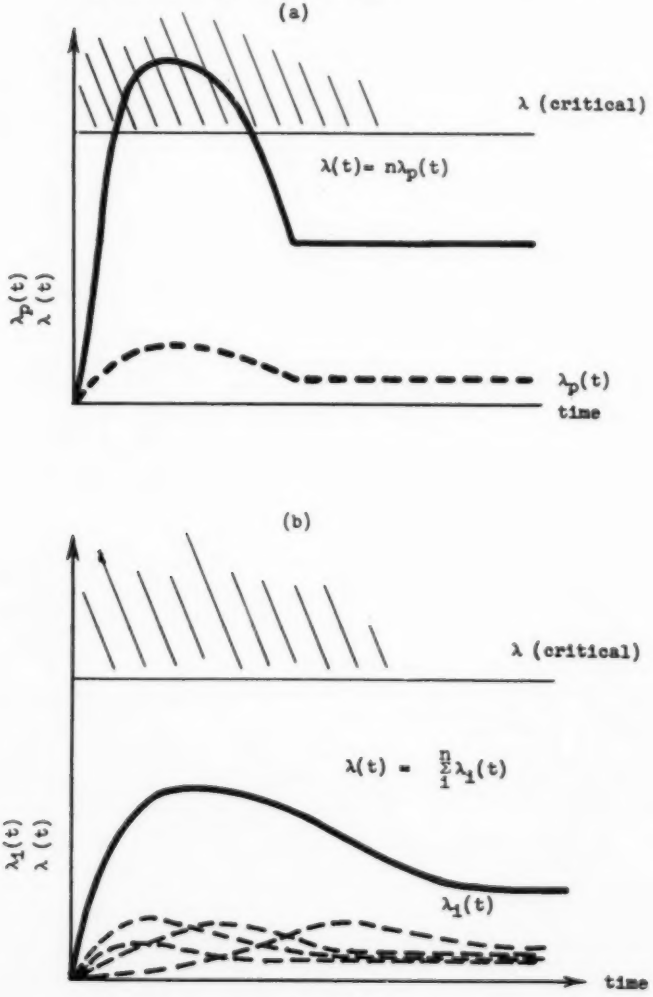


Figure 3

3. Thirdly, if a part is to be "burnt-in" prior to use, the required "burn-in" period cannot really be adequately established without constructing the complete time dependent λ characteristic. This can be seen readily in Figure 4. If the system operating period is t_1, for example, a "burn-in" period is unnecessary and even harmful since $\lambda(t_1) < \lambda(t_4)$. If the system operating period is t_2 or greater, a "burn-in" period is clearly desirable.

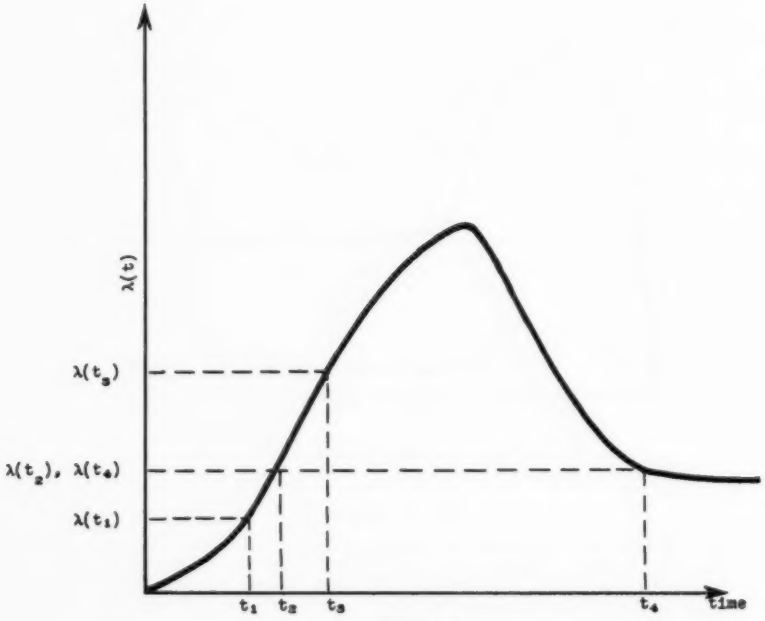


Figure 4

The last three figures, and some of the associated arguments, have been adapted from Druzhinin's article(3) in the book "Reliability of Radio-Electronic Apparatus," published in 1958 by "Soviet Radio." This book, incidentally, is the first of promised annual publications of collections of research papers in the field. Apparently in this, as in so many other fields, U.S.S.R. technical organizations have initiated a systematic, broad-based approach. The National Bureau of Standards, in general, and Joan Rosenblatt, in particular, are to be thanked for their initiative in providing translations of some of the more important Russian papers in the reliability area.

Having established λ characteristics for all parts in the system, we can then directly employ the results of the previous analysis for systematic construction of the reliability functions for components, subsystems, and the over-all system.

The procedure can be illustrated by the argument shown in Figure 5, where $R_p(t)$ is the "part" reliability function.

$$R_p(t) = e^{-\int_0^t \lambda_p(t)dt}$$

For a component of n "independent" parts:

$$\lambda_c(t) = \sum_{i=1}^{n} \lambda_i(t) \text{ , and}$$

$$R_c(t) = e^{-\int_0^t \sum_{i=1}^{n} \lambda_i(t)dt}$$

- - - - - - - - - - - - - - - - -

If there are K identical parts in a simple redundant arrangement, the groups reliability function is:

$$R_g(t) = 1-\left[1-R_p(t)\right]^k \text{ ,}$$

where $R_p(t)$ has the form shown above.

The component reliability function, $R_c(t)$, /assuming (n-k) "independent" parts and a single group of k parts in a redundant arrangement/ is then:

$$R_c(t) = R_g(t)e^{-\int_0^t \left(\sum_{i=1}^{n-k}\right)\lambda_i(t)dt}$$

Figure 5

If a component, for example, consists of n parts, which our analysis has shown to have independent failure probabilities, then the λ characteristic for it is simply the sum of the part characteristics and its reliability function is as shown in Figure 5. In the general case, $\lambda_c(t)$ must be obtained by the detailed superposition process of the type previously shown in Figure 3.

On the other hand, if the component has been found, through our analysis, to have a group of parts which must be treated as an entity with respect to independence of failure probability in relation to the other parts, the reliability function for the group must be built up in accord with the logical relations found for the parts in the group. Probably the simplest case of this sort occurs when the group's parts merely provide functional redundancy. In such a case, the group and component reliability functions can be obtained as is shown in Figure 5.

The argument for other component, for the subsystems, and the over-all system then proceeds in an analagous way.

Here it should also be mentioned that another important by-product of the general method outlined is that if the resultant over-all system reliability is found to be, for example, unsatisfactorily low, a firm basis has been established already for evaluating the regions of the system where increased part or component reliability, or the employment of redundancy, will be most effective in raising the over-all reliability of the system.

In discussing part testing as a part of the job of predicting the reliability of systems in early stages of development, it was implied that the parts to be used in the system may not actually be available for tests to determine their λ characteristics. Of course, if such restrictions exist, there is little choice but to use available failure rate information for similar parts previously tested or used under closely related environmental conditions. However, this is an extremely dangerous procedure, at least for systems whose development cycles extend over several years. This is the case because technology is advancing so rapidly in some fields that errors of several orders of magnitude are possible unless careful and explicit allowance is made for the changing state-of-the-art. Figure 6, taken from a recent IBM report(4), shows what is expected to happen in a field of major importance to military applications -- that of computers.

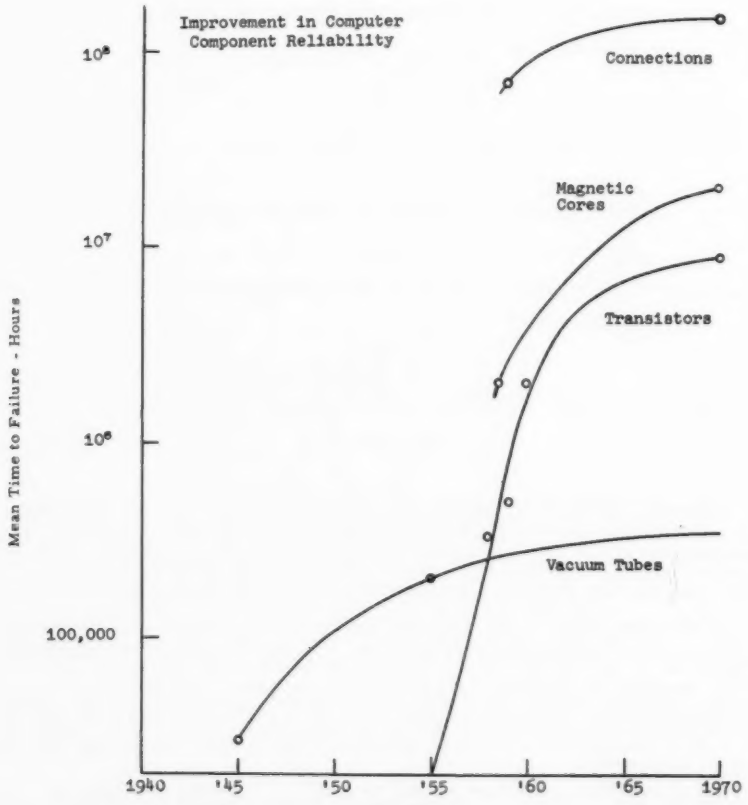


Figure 6

Looking back over what has been said, there are two points which have not been touched on; namely, the hierarchical status of reliability responsibility and the matter of whose business, in general, is reliability assessment and prediction. Although these aspects, especially when voiced explicitly, appear fatuous, they have certain nontrivial implications. In our experience, for example, we have obtained, from highly responsible and competent groups, estimates of the reliability of a proposed complex system which differ by a factor of more than 10. A few further remarks may, therefore, be in order.

1. In the first place, a key requirement in our point of view is clear articulation of the system design into associated mathematical structures and computer-simulation schemes. Such an undertaking must necessarily be assumed either by the design staff itself or by a group otherwise living with the job. The requirement for this association is not just a matter of the complexity of the task, but also that the design program itself will benefit enormously through numerous feedbacks from a thoroughgoing application of the suggested procedure.

2. In the second place, adequate resources must be provided for part development, procurement, and testing; sufficiently realistic environmental conditions must be available for the program, and means must exist for taking advantage both of information available on part and component performance from other contexts, and of changing technology which rapidly makes most extant information quickly obsolete. These tasks, most conveniently, must also be closely associated with the engineering group responsible for design.

3. In the third place, strong motives must exist for a realistic approach to the problem of reliability prediction if useful results are to be made available. This is clearly the case because of the frequency of situations, particularly in the case of complex systems, where no useful data are available and judgments must be substituted. Strong motivations to conservative and realistic prediction obviously can be found only if financial losses rather than gains are to be expected from lack of realism. These considerations suggest that the guidance and detailed technical review necessary for an adequate and realistic reliability prediction program cannot usually be expected to be found ready-made in the group responsible for justifying the feasibility and usefulness of a design. This is particularly the case with complex systems for which it is out of the question to require a serial post-review of the designer's estimates -- simply because the task is too big and will probably require more time than is reasonable to allow in postponing decisions to accept or not accept a given design.

There is, of course, no question that the system design contractor must have a competent reliability analysis group and that, at least administratively, it should not be under the direct management control of the design group itself.

However, the above remarks suggest, further, particularly in the case of government procurement actions for new complex systems in which matters of operational reliability are of basic importance, that somewhat more attention than may have been customary in the past be given to developing reliability analysis and prediction programs coincident in time with the beginning of a system design. As a matter of fact, it can probably be persuasively argued that a thoroughgoing, coincident effort is likely to be not only more fruitful but also, in the final summing up, including operational phases, much less costly.

Also, it seems fairly clear that such reliability analysis and prediction programs should proceed either under the direct, in-house guidance of the

government or be conducted under such guidance assisted by adequately motivated contractors not themselves committed to major R&D management or hardware programs in the systems being so studied.

REFERENCES:

(1) MIL-STD-441, Reliability of Military Electronic Equipment, 20 June 1958, Office of the Secretary of Defense, Supply and Logistics.

(2) Exploratory Conference on Missile Model Design for Reliability Prediction, Report of the Third Meeting, White Sands Missile Range, New Mexico, 23 April 1959.

(3) G. B. Druzhinin, On Methods of Calculating the Reliability of Systems, Soviet Radio, Moscow, 1958.

(4) Digital Computer Characteristics for Space Applications, IBM Corporation, Federal Systems Division, Oswego, New York, June 9, 1959.

RELIABILITY ASSURANCE PROVISIONS IN SPECIFICATIONS

Charles J. Brzezinski
Office of the Assistant Secretary of Defense
(Supply and Logistics)

SUMMARY

Experience has shown that component parts produced to existing specifications are often quite reliable. However, not all parts produced to the same specification are equally reliable. Some are capable of meeting only the minimum quality and reliability level established by the specification as a sole and inflexible standard. Thus, if a part is purchased to an existing specification, there is no assurance that its reliability will be any better than this standard minimum level.

The procedure described in this paper proposes that specified test procedures should be revised to eliminate existing minimum quality and reliability levels as standards. A new approach is proposed which will provide quality assurance provisions that (1) establish more than one standard reliability level for each product and (2) provide an incentive for product improvement through a qualification procedure which identifies each product with the reliability level it has attained.

INTRODUCTION

The purpose of this paper is two-fold: (1) to identify those specification requirements which determine whether a component part will meet a desired level of reliability and (2) to illustrate a practical method for incorporating reliability assurance provisions into specifications for component parts.

The primary purpose of this paper is to illustrate a method for revising certain specification requirements which in their present form are incapable of distinguishing between the most reliable and least reliable parts produced to the same specification. It is common knowledge that many component parts far exceed the minimum quality and reliability levels established as standards by existing specifications. However, existing specification requirements are not designed either to identify these more reliable parts or to indicate the level of reliability they can meet. This paper describes a new approach to test procedures in specifications whereby the purchaser of a component part can be assured that product does meet a desired reliability level.

At the outset it might be well to explain the circumstances that motivated this paper. These circumstances pertain primarily to problems that confront quality control and reliability engineers who have the responsibility for reviewing and approving, and possibly improving, specifications prior to their publication. Not only must the provisions of new specifications be carefully evaluated but, also, existing specifications must be reviewed to assure that they are responsive to current needs.

In view of the present emphasis on reliability, it is understandable that engineers responsible for reviewing specifications inevitably ask themselves two questions, among others: (1) Which requirements in this specification affect reliability? (2) How effectively does this specification satisfy new or anticipated reliability requirements? One's response to each of these questions depends, in part, on where one stands. A research and development engineer, for example, might evaluate a proposed specification from a somewhat different point of view than a production or quality control engineer.

This paper is written from the point of view of the quality control engineer. The question which he seeks to answer is whether the specification is capable of assuring the desired degree of reliability. How well he is able to answer this question may determine whether the product will perform satisfactorily as a component in the assemblies of which it becomes a part.

To answer the aforementioned questions, the engineer must first identify those provisions of a specification that directly affect reliability. Having done so, he is in a position to suggest such improvements as may be necessary to assure adequate reliability.

Examinations and Tests

For an evaluation of the effectiveness of existing specification provisions, the following procedure is both simple and practicable: (1) Establish basic groups of tests, with each group representing a clear-cut explicit function. (2) Review each existing test and then classify that test into its appropriate functional group.

It will be found that tests as set forth in existing specifications fit into one of the following three groups:

(1) Group 1, Conformance Inspections - These inspections are performed to ascertain that product conforms to standards established by the design. Each examination or test in this group involves a measurement or comparison to a design standard.

(2) Group 2, Environmental Tests - These tests are performed to ascertain whether product conforming to design standards (per Group 1 above) can withstand certain environmental conditions to which it may be subjected in operation. These tests include environmental conditioning (e.g., subjection to shock, vibration, temperature variation, etc.) during or after which Group 1 (Conformance Inspections) evaluations are repeated to detect any change in the product.

(3) Group 3, Life Tests - These tests are intended to ascertain whether product conforming to design standards is capable of functioning satisfactorily throughout its projected life. These tests are characterized by a specified number of hours or cycles of functional operation during or after which certain Group 1 (Conformance Inspections) evaluations are repeated to detect any changes in the product.

When the examinations and tests which have a common purpose are grouped together, it is possible to determine whether suitable inspection procedures have been specified for the tests in each group. The following principles are helpful in determining the type of procedure that should be specified:

(1) Conformance inspections should be planned to detect defects which may be acquired at any time during the manufacturing process.

(2) Environmental tests should be planned to detect design errors. These errors may be generated before production begins and at any time a design change is made.

(3) Life tests are intended to detect limited life potential caused by certain combinations of normal variations in design and construction detail. Undesirable variations may be generated either by design error or by the manufacturing process.

Typical Specification Requirements

Using the above criteria as a guide, let us now examine the content of a specification for a typical component part. This examination shows that technical requirements for the part are normally expressed in terms of performance characteristics. Those few material, design and construction details which are specifically described define little more than the "envelope" characteristics of the part. Within this envelope each manufacturer of the part is permitted to choose whatever details of material, design and construction he determines to be most suitable.

Such a performance-type specification is both realistic and practical for any but the simplest kind of product. If there are many possible combinations of material, design and construction detail that might be used, the part manufacturer and not the purchaser is most qualified to decide which combination of these details will produce the best product. In fact, the search for a better combination is continuous and, unless there are certain overriding considerations, it is undesirable to "freeze" design detail and thereby prevent improvements.

The performance-type specification should, of course, embody quality assurance provisions which do not have the effect of limiting product improvement. Customarily, however, specifications tend to establish a minimum quality level as the standard. This discourages the part manufacturer from initiating any product improvement that might increase costs and, in effect, establishes an incentive to produce the cheapest part which can meet the minimum quality standard.

RELIABILITY ASSURANCE

The quality engineer's problem, then, is to assure that the quality assurance provisions of a specification do not establish an undesirable minimum quality and reliability level as the standard. In fact, his problem is to assure that no minimum reliability and quality level is established as a sole and inflexible standard. Obviously, some minimum level must be established but, if possible, provisions should be made for automatically adjusting this minimum level to whatever higher level the product can attain.

Such flexible quality and reliability levels are a radical departure from the present concept of quality assurance provisions in specifications. It is, however, the only practical means for obtaining the most reliable components with the least amount of testing. If, as appears to be the case, parts are being produced that exceed the minimum quality and reliability levels established by existing specifications, then the only problem is to identify these more reliable parts. This can be done, and is being done, by performing more testing than now called for by the specifications.

If effect, the establishment of flexible quality and reliability levels in specifications involves the development of test procedures that will automatically adjust the amount of testing to the level a product can meet. This can be done by selecting certain tests applicable to the product and specifying a suitable sampling plan for these tests.

Life Testing for Reliability Assurance

Referring to the three groups of examinations and tests described above, it will be noted that only one group, namely, Life Tests, can be used to estimate product reliability. Properly planned, the life tests can provide an estimate of the capability of a product to perform its intended function throughout its expected life. Thus, depending on the sampling procedure specified, the life test can be used to assure product conformance to a desired reliability level.

To be effective, the life test requirement should be inherently flexible to provide not only for evaluation of the reliability of units of current production, but also for the evaluation of new reliability levels attained if product reliability improves. Specifically, the life tests must be performed on samples that are truly representative of production under evaluation and are large enough to make the life test meaningful. This poses a rather difficult problem in some instances because life tests have two undesirable characteristics. First, life tests are destructive; secondly, either the life test must be lengthy or the sample tested must be large. Thus, the development of a suitable life test procedure taxes the ingenuity of both engineer and statistician.

With respect to the problem of large sample sizes and lengthy tests, the use of complex and sophisticated statistical techniques provides only a partial solution. A more effective approach is to develop accelerated life testing methods. Referring to the example in Appendix B, it will be noted that the life test used is an accelerated test. Although the acceleration factor is unknown, it is possible to use such a test if, as in this example, it is not necessary to describe the reliability level in the form of a numerical expression.

This is called to the readers' attention because there may be many instances where accelerated life tests must be used without specific knowledge of the degree to which product life is accelerated. If it is desired to establish reliability levels that are significantly different one from another, it is not necessary to know how much life is accelerated. On the other hand, if it is important to obtain an accurate estimate of the failure rate, life testing procedures must be used that can be converted to some standard condition.

Sampling for Reliability Assurance

Since life tests are destructive and cannot be performed on the parts that will be used in equipment, the reliability of any part can only be estimated. In fact, until a great amount of life testing has been performed, the estimate may be quite inaccurate because it must be based on an assumption. If, for example, the reliability of a product were expressed as failure rate in terms of "Y% failures per 1000 cycles of operation," the figure for Y would have the following meaning:

"Assuming that the distribution of time-to-failure conforms to a certain statistical pattern when a sample is subjected to the specified life test, there is an X% probability that, if no more than a certain number of failures occur in the sample, the failure rate of the parts tested does not exceed Y% per 1000 cycles of operation."

A failure rate figure, Y% derived in accordance with the above definition obviously has a very limited meaning. Until a great amount of testing has been completed, the distribution pattern of time-to-failure is unknown. It is possible only, by the choice of a suitable sampling plan, to control the X% probability figure. This fact must be made clear to design engineers: It is impossible to derive a meaningful numerical value for reliability from a life test that is not performed under actual use conditions.

For these reasons, no numerical values are shown for the reliability levels established in the example in Appendix B. As shown in Appendix A, reliability levels have been established by assuming an exponential distribution of time-to-failure and by use of sampling procedures that will reject 90% of the time any sample worse than the established level. The amount of testing is, of course, significantly greater than required by the existing specification, but very much less than any user of the part must now perform to obtain equivalent assurance of product reliability.

Qualified Products List (QPL)

Product Qualification is, of course, an essential element of a multi-level quality and reliability assurance provision. Unless there is evidence that a product has attained a specified reliability level, there is no certainty that it ever will. As a practical measure, therefore, evidence that a product can meet the desired reliability level should be obtained before a contract is placed.

The present Qualification Inspection program and the QPL are used for essentially the same purpose. Thus, the existing qualification procedure can be used to control and identify those products which meet the various reliability levels established in specifications. Referring to the example in Appendix B, it will be noted that the basic qualification procedure has not been changed other than to retitle it as "Initial Qualification." This was necessary to distinguish the initial qualification from subsequent Reliability Level Qualification.

CONCLUSION

A component part purchased to an existing specification may or may not be reliable because existing test procedures recognize conformance to only one minimum level of reliability. If these test procedures are suitably revised, establishing more than one reliability level, purchasers of the component will be able to clearly identify the reliability level desired. With respect to component parts, a major part of the reliability problem can be solved by specifying test procedures that can distinguish between products that meet significantly different levels of reliability.

REFERENCES

E. C. Molina, "Poissons' Exponential Binomial Limit," D. Van Norstrand Co., Inc., New York 1947

National Bureau of Standards, mathematical tables, MT2, "Tables of the Exponential Function e^x," 1939

A. Wald, "Sequential Analysis," John Wiley & Sons, 1948

I. B. Altman, "Sampling Plans for Determining Conformance with Electronic Parts Reliability Requirements," Office of the Assistant Secretary of Defense (Supply & Logistics)

Appendix A

Problem

The life test (current overload) sampling plan in the "first quantities produced" requirement of MIL-C-3737C calls for a sample of 6 and acceptance number 0. Assuming that this sampling plan establishes a minimum reliability level for connectors produced to the specification, derive the following:

(1) A lot-by-lot life test sampling procedure for this minimum reliability level, Level A, such that there is a 90% probability of rejecting any lot that exceeds this level.

(2) A similar sampling plan providing equivalent protection for a second level, Level B, which is only sufficiently higher than Level A for a statistical determination that the two levels are different.

(3) A sequential analysis sampling procedure, based on the sample sizes specified for Level B, for life testing at Levels C, D and E. The sequential sampling plan is to provide protection equivalent to that of an LTPD type plan for these levels. The difference between levels is to be no greater than necessary.

(4) Qualification and disqualification criteria for each level.

The following code is used in the mathematical computations:

c - Acceptance number

n - Number of units in sample
N - Total number of sample units tested

p_1 - Fraction defective expected in sample at P_1
P_1 - .10 probability of occurrence

p_o - Fraction defective expected in sample at P_o
P_o - .95 probability of occurrence

t - Length of life test in terms of cycles of operation
T - Accumulated total number of test cycles

F - Accumulated total number of failures

R - Failure rate as a fraction per 1000 cycles of operation

Solution of Problem

(1) Sampling Plan for Reliability Level A:

Given n=6 and c=0 when t=50 cycles

*From tables of Poisson's Exponential Binomial Limit:

$np_1=2.3$ (when $P_1=.10$ and c=0 and $P_1=np_1/6=.383$

Assuming an exponential distribution of time-to-failure:

$e^{-\frac{Rt}{100}}=1-p_1=.617$ and
$R=.9658$

Thus, when R=.9658 and t becomes 500 cycles,

$(1-p_1)=.00745$ and
$P_1=.99265=.993$

Using $p_1=.993$ compute sample sizes for c=0, 1, 2, 3 and 4:

c	np_1	P_1	np_1/p_1	n
0	2.3	.993	2.3	2
1	3.9	.993	3.9	4
2	5.3	.993	5.3	5
3	6.7	.993	6.8	7
4	8.0	.993	8.1	8

(2) Sampling Plan for Reliability Level B:

In the sampling plans computed for Level A, p_o is highest when n=8 and c=4

Let p_o at Level A=p_1 at Level B

From the tables $np_o=1.97$ and $p_o(A)=1.97/8=.246=p_1(B)$

c	np_1	$p_1(B)$	np_1/p_1	n
0	2.3	.246	9.3	9
1	3.9	.246	15.8	16
2	5.3	.246	21.5	22
3	6.7	.246	27.2	27
4	8.0	.246	32.4	32

(3) Sequential Analysis Sampling, Reliability Level C:

*Note - In order to facilitate explanation Poisson tables are used rather than Binomial tables.

Since Level B sample sizes are used for Level C, there is a 95% probability that a C Level lot will pass the test. Therefore, if a sample passes the test, the lot is not code marked Level C until an analysis is made of the life test data from a specified number of successive lots. The following indicates the degree of protection attained:

a. 95% probability that a sample will pass the life test if relays in the lot meet Level C.

b. 90% probability that if a sample passes the life test, the lot is not less than Level B.

c. 90% probability that a lot will not be code marked Level C if the average level for a specified number of consecutive lots is below Level C.

Using the largest sample size for Level B (n=32, c=4)

P_o at Level B=np_o/n=1.97/32=.0615=p_1 at Level C

When c=4, np_1 at Level C=8.0 and n=8.0/.0615=130

Then np_o at Level C=1.97 and P_o=1.97/130=.0151

Compute the "reject" line for a sequential analysis such that there will be:

1. A .05 probability of rejecting when lot reliability = p_o
2. A .90 probability of rejecting when lot reliability = p_1

The equation for this reject line is F=a+bN

Since T=Nt and t=500 F=a+bT/500 When:

$$a=\frac{\ln(1-p_1)-\ln(1-P_o)}{\ln p_1-\ln p_o+\ln(1-p_o)-\ln(1-p_1)}$$

$$b=\frac{\ln(1-p_o)-\ln(1-p_1)}{\ln p_1-\ln p_o+\ln(1-p_o)-\ln(1-p_1)}$$

Substituting for P_1, P_o, p_1 and p_o,

$$a=\frac{\ln.90-\ln.05}{\ln.0615-\ln.0151+\ln.9849-\ln.9385}=1.99$$

$$b=\frac{\ln.9849-\ln.9385}{\ln.0615-\ln.0151+\ln.9849-\ln.9385}=.033$$

Since F=a+bT/500

F=1.99+.033T/500=1.99+.000 066 T

=2.0+6.6 x 10^{-5}T

(4) Qualification Criteria, Reliability Levels B through E:

For practical reasons, if a product is qualified for a reliability level, it should not be subject to sudden and unexpected disqualification. Due to sampling error, such disqualification may occur unless suitable control limits are established. Under the procedure described in this paper, the widest possible limits can be specified for two reasons: (a) the life test sampling procedure will reject 90% of lots which exceed the reliability level and (b) qualification for a higher level can be obtained only if a sequence of life tests indicates a 98% probability that product meets the higher level.

Control limits for fraction failures,

$$p=p_1 \pm 2\sqrt{p_1(1-p_1)/N} \qquad (1)$$

Since p=F/N and N=T/t then p=Ft/T

Substituting in (1):

$$Ft/T=p_1 \pm 2\sqrt{p_1(1-p_1)t/T}$$

and $$F=p_1T/t-2\sqrt{p_1(1-p_1)T/t}$$

For Level B, p_1=.246 and t=500

Thus, the maximum number of failures permitted to qualify for Level B:

$$F=.246T/500-2\sqrt{.246x.754xT/500} \qquad (2)$$

And the maximum number of failures permitted to retain qualification:

$$F=.246T/500+2\sqrt{.246x.754xT/500} \qquad (3)$$

F in (2) above is a minus quantity until T=the minimum number of hours required for qualification which is when F=0 and (2) becomes:

$$0=.246T/500-2\sqrt{.246x.754xT/500}$$

And T=6160 cycles.

Appendix B

(This is intended as an example of a specification revised to embody reliability requirements and has not been approved by the DoD for use by the Departments of the Army, Navy and Air Force.)

MIL-C-3767A
CONNECTORS, ELECTRICAL (POWER, BLADED TYPE), GENERAL SPECIFICATION FOR

1. SCOPE

1.1 Scope. This specification covers the general requirements for power-type electrical connectors (plugs and receptacles) with bladed contacts, for use in direct-current (dc) circuits up to 600 volts, or in alternating-current circuits up to 600 volts at frequencies up to 400 cycles.

1.2 Classification

1.2.1 Type designation. The type designation shall be in the following form and as specified (see 3.1 & 6.2):

* * * * *

1.2.1.3 Reliability level. The reliability level for which a connector is qualified is identified by a single letter. The lowest reliability level established by the letter "A." Higher levels are identified by letters selected in alphabetical order.

* * * * *

3.1.1 Qualification. Connectors furnished under this specification shall be a product which has been tested and has passed the Initial Qualification and Reliability Level Qualification requirements specified in Section 4. (See 6.1)

* * * * *

4.1 Classification of inspection. The examination and testing of relays shall be classified as follows:

(a) Conformance inspections
(b) Environmental tests
(c) Life tests

4.1.1 Responsibility for inspection. Manufacturers are responsible for the performance of all inspections specified herein. Except as otherwise specified, manufacturers may use their own or any other laboratory facilities acceptable to the procuring activity.

4.2 Qualification.

4.2.1 Initial qualification. A sample of connectors shall be subjected to the Qualification tests listed in Table I. Table I shows the number of specimens required and the order in which the tests are to be performed.

4.2.1.1 Defectives. Defectives in excess of those allowed in Table I, subgroups 2 through 7, will be cause for refusal to grant initial qualification.

4.2.2 Reliability level qualification. Initial qualification includes qualification for the lowest reliability level permitted by this specification. Qualification at higher levels shall be based on accumulated life test data. Both normal and extended life test data shall be used in determining whether a connector qualifies for a higher reliability level.

4.2.2.1 Qualification for higher levels. Qualification for a higher reliability level shall be granted by the qualifying agency upon submission by the manufacturer of sufficient life test data to prove that connectors from current production consistently meet the reliability level for which qualification is desired.

4.2.3 Qualification re-evaluation. Subsequent to qualification, the connector shall be subject to periodic re-evaluation by the qualifying agency to assure continued ability of units from current production to meet the qualification requirements. Qualification may be revoked under either of the following conditions:

(a) Periodic environmental tests indicate that units from current production may be incapable of meeting qualification inspection requirements.

Table I - Applicable Examinations and Tests

Test	Requirement Paragraph	Method Paragraph	No. of Specimens	Allowable Failures
Conformance Inspection				
Subgroup 1				
Visual & mechanical (Internal & external.....	3.1, 3.3 to 3.3.3.2 incl. 3.3.5 to 3.4.5 incl.	4.6.1		
Strain relief.....	3.5	4.4.2		
Dielectric withstanding voltage.....	3.6	4.4.3		
Insulation resistance.....	3.7	4.4.4		
Coupling & uncoupling force.....	3.8	4.4.5		
Voltage drop.....	3.9	4.4.6		
Contact strength.....	3.10	4.4.7 to 4.4.7.2 incl.		
Qualification Tests - Environmental & Life				
Subgroup 2 Tests				
Terminal strength.....	3.11	4.4.8		
Temperature cycling.....	3.12	4.4.9	6	0
Moisture resistance.....	3.14	4.4.11		
Subgroup 3				
Compression.....	3.15	4.4.12		
Vibration.....	3.16	4.4.13		
Drop test.....	3.17	4.4.14	6	0
Salt spray (Corrosion)....	3.18	4.4.15		
Heat resistance.....	3.19	4.4.16		
Subgroup 4				
Tensile strength.....	3.3.4.1	4.4.18	6	0
Ultimate elongation.....	3.3.4.2	4.4.19		
Subgroup 5				
Air-heat resistance.....	3.3.4.3	4.4.20	3	0
Subgroup 6				
Oxygen-pressure resistance.....	3.3.4.4	4.4.21	3	0
Subgroup 7				
Life test.....	3.13	4.4.10	6	0

(b) Life test data shows that the item has not attained a reliability level equal to or higher than the lowest reliability level specified herein.

4.2.3.1 Periodic environmental tests. The environmental tests specified in Test Groups 2 through 6 of Table I shall be performed in the order shown on a sample selected in accordance with 4.2.3.2.

4.2.3.2 Sampling procedure. 12 connectors and 12 rubber specimens shall be taken from production every 3 months for whichever type of connector is being manufactured during the week when the sample is selected.

4.2.3.3 Disposition of sample units. Connectors subjected to periodic environmental tests shall not be offered for delivery.

4.2.3.4 Environmental test failure. If a sample fails to pass periodic qualification re-evaluation tests, action shall be taken immediately to determine the cause of failure. Connectors represented by the sample and all other connectors manufactured to the same design, which have not been delivered, shall be held until the cause of failure has been determined. If it is determined and concurred in by

Table II - Sampling for Life Test

Acceptance Number	Sample Size Corresponding to Reliability Level				
	A	B	C	D	E
0	2	9	(See 4.4.3)		
1	4	15			
2	5	21			
3	7	27			
4	8	32			

the qualifying agency that the failure is not cause for disqualification, the connectors being held may be offered for delivery.

4.3 Conformance inspection. Conformance inspection shall consist of Test Group 1 in Table I.

4.3.1 Point of inspection. Connectors shall be inspected to assure conformance to all material, design and construction details established by the design. Conformance inspection shall be performed at whatever point in the manufacturing process is deemed most suitable.

4.3.2 Conformance inspection sampling. If conformance inspection is performed on a sampling basis, a suitable sampling plan shall be used, capable of protecting the reliability level for which the product has qualified.

4.4 Reliability testing. Reliability testing shall consist of the life test (Requirement 3.13, Test Method 4.4.10.2).

4.4.1 Selection of sample. A sample shall be selected at random from each inspection lot which has passed conformance inspection. The manufacturer may use any sample size and corresponding acceptance number shown in Table II, provided (a) the sample size is selected before the life test begins and (b) the sample size selected is in the group of sample sizes listed under the code letter corresponding to the reliability level for which the connector is qualified.

4.4.2 Inspection lot. An inspection lot shall consist of all connectors of the same type designation, manufactured as a continuous production run, or at the option of the manufacturer, all connectors of the same type designation comprising one week's production. Each lot shall be kept separate from every other lot, and shall be identified in a suitable manner to prevent the mixing of connectors from one lot with connectors from another lot.

4.4.3 Sequential analysis sampling procedure. If a product attains a reliability level higher than B, a sequential analysis sampling procedure shall be used for life tests. Samples for life tests shall be selected in accordance with 4.4.1 for any level higher than B.

4.4.3.1 Individual lot acceptance or rejection. Under the sequential analysis procedure, individual lots shall be rejected or accepted in accordance with the sample size and corresponding acceptance number selected for the life test.

4.4.3.2 Rejected lots. For connectors qualified at the A and B levels, failure of a sample to pass the life test shall result in final rejection of the lot. For connectors qualified at the C level or higher, a lot that fails to pass the life test may, at the option of the manufacturer, be rejected or marked as conforming to the A level.

4.4.3.3 Reliability level identification. For connectors qualified at the A or B level, if the sample passes the life test, all connectors in the lot shall be marked with the code letter corresponding to the level for which the connector has qualified. For connectors qualified at the C or higher level, if a sample passes the life test, appropriate code marking shall be determined in accordance with 4.4.3.5.

4.4.3.4 Tabulation of life test data. Under the sequential analysis sampling procedure, life test data shall be tabulated in the order in which tests are completed on a form similar to Figure 1, including data from both accepted and rejected lots.

4.4.3.5 Reliability level C and higher. For connectors qualified at the C level or higher, code marking shall be determined on the basis of

Figure 1

Summary Report of Life Test Data

Lot Code	Lot Size	Acceptance Life Test			Extended Life Test			Lot Totals	
		Sample	Total Test Cycles	No. of Failures	No. Tested	Total Test Cycles	No. of Failures	Total Cycles	Total Failures
							TOTALS		

summarized data from at least the last 10, but not more than the last 15 lots for which life tests have been completed. As each new entry is made in the tabulation required in 4.4.3.4, the last 10 to 15 entries shall be summarized and the permissible number of failures computed by using the appropriate equation in Table III. If the actual number of failures in the lots used in the tabulation does not exceed this computed figure, the last lot tested will be identified as meeting the reliability level for which the product is qualified. However, if the computed figure is exceeded, the last lot will be identified as meeting the B level.

Table III - Equations for Computing Permissible Number of Failures

Reliability Level	Permissible Failures, F, for Total Test Hours, T
C	$F = 2.0+6.6\times10^{-5}T$
D	$F = 2.0+2.2\times10^{-6}T$
E	$F = 2.1+5.4\times10^{-7}T$

4.5 Reliability Level Qualification Procedures

4.5.1 Summary of life test data. The manufacturer shall summarize life test data on a form similar to Figure 1. The summary shall be prepared as of the last day of each calendar month. Test data shall be entered on the form in the chronological order in which tests were performed.

4.5.1.1 Amount of data in summary. Table IV shows the minimum number of test operations required for reliability analysis purposes. If the amount of life testing during any monthly summary period is insufficient, data from the preceding period shall be included in the summary.

4.5.2 Qualification for higher reliability levels. Qualification at the next higher level may be applied for when the number of failures during a summary period is below the maximum number of failures permitted by the appropriate equation in Table IV.

4.5.2.1 Retention of qualification. Qualification at a reliability level shall be retained if the number of failures during any summary period does not exceed the maximum number permitted by Table IV.

4.5.2.2 Revocation of qualification. If the number of failures during any summary period exceeds the maximum number permitted by Table IV, qualification will revert to the next lower reliability level.

4.5.3 Additional life test data. Where normal life testing operations do not provide sufficient data to determine qualification for a higher level, the manufacturer is permitted to subject not more than 10% of the units in the sample to the extended life test, provided the units subjected to the extended life test are selected before the test begins.

* * * * *

4.6.10 Current overload. The female connector and the applicable test plug shall be coupled and wired as for actual service with suitable conductors. The rated voltage (rms) $\pm$ 5 percent and 150 percent of the maximum rated current (see 3.1) shall be applied. The power factor of the load shall be within the limits of 0.75 and 0.80. The test plug shall be alternately coupled and uncoupled from the female connector for 50 complete cycles of operation, at a rate not exceeding 10 cycles per minute. At the completion of the last cycle, the voltage drop dielectric with-

Table IV - Control Limits for Qualification at Each Reliability Level

Reliability Level	Minimum Number of Accumulated Test Cycles in Summary	Maximum Number of Failures Permitted During Summary Period to Qualify T = Total Test Cycles	Maximum Number of Failures Permitted During Summary Period to Retain Qualification T = Total Test Cycles
A			
B	6,100	$F = .0005T - .04\sqrt{T}$	$F = .0005T + .04\sqrt{T}$
C	30,400	$F = .00012T - .02\sqrt{T}$	$F = .00012T + .02\sqrt{T}$
D	131,000	$F = .00003T - .01\sqrt{T}$	$F = .00003T + .01\sqrt{T}$
E	536,000	$F = .000007T - .005\sqrt{T}$	$F = .000007T + .005\sqrt{T}$

standing voltage, and coupling and uncoupling force shall be determined in accordance with 4.4.6, 4.4.3 and 4.4.5.

4.6.10.1 Initial qualification. For initial qualification, the life test shall be performed in accordance with 4.6.10.

4.6.10.2 Reliability tests. Life tests shall be performed in accordance with 4.6.10 except that the test length shall be 500 cycles, and required measurements shall be made at the end of each 100 cycles of operation.

4.6.10.3 Extended life test. Connectors subjected to reliability tests may be subjected to an additional 4500 cycles of operation in accordance with 4.5.3 and 4.6.10.2.

* * * * * * * *

6.1 Qualification

6.1.1 Qualification. With respect to products requiring qualification, awards will be made only for such products as have, prior to the bid-opening date, been inspected and qualified for inclusion in the applicable Qualified Products List whether or not such products have actually been so listed by that date.

6.1.2. Reliability level qualification. Where the invitation for bid specifies a reliability level above the minimum level permitted by this specification, awards will be made only for such products as have, prior to the bid-opening date, been qualified for the reliability level specified in the invitation for bids.

POWER CHARACTERISTICS OF CONTROL CHARTS*

Edwin G. Olds
Carnegie Institute of Technology

*Presented, by invitation of the Program Committee, at the Fourteenth Annual Convention of the American Society for Quality Control, at San Francisco, California, May 25, 1960.

Prepared in connection with research sponsored by the Office of Ordnance Research, U. S. Army. It is released from copyright restrictions.

INTRODUCTION

This paper will present a brief summary of results obtained by the author and several of his graduate students at Carnegie Institute of Technology during the past decade. Since September, 1954 the research has been sponsored by the Office of Ordnance Research, U. S. Army.

The application of statistical quality control has two phases. The first phase is described [12]** as "control when no standard is given" and the second

**Bracketed numbers refer to items in the list of references at the close of this paper.

phase as "control with respect to a given standard." The operating characteristics of the average and range charts, in the second phase for the case when the measured quality characteristic has a Gaussian (normal) distribution, were given in a fundamental paper [16] by Scheffe in 1949. The investigations to be discussed here have been concerned, mainly, with power characteristics of the control test in the first phase - the phase where the control chart is, in the words of $z_{1.3}$ [2], used "for analyzing past data." Control charts for both counts and measurements are considered but, in the latter case, results can be given only when the assumption of normality is imposed.

Actually, the recognition of the need for the investigation of power characteristics for the first phase antedated the "fabulous fifties" by at least another ten years. It was strongly emphasized by the embarrassment of the author when he was associated with the Office of Production Research and Development during the last war and was faced with the problem of explaining why the control chart works. He was forced to restrict himself to the making of definite statements about phase two because specific information regarding phase one was almost completely lacking.

An incident at one of the intensive eight-day courses, organized by the Office of Production Research and Development, was particularly disconcerting. (For information regarding the intensive courses see [20] and [13]). Early in this course on "Quality Control by Statistical Methods," the behavior of samples from a controlled process was simulated by the drawing of twenty-five samples of five chips from the same bowl. As usual, averages and ranges were calculated and plotted. Then control limits were calculated and placed on the charts. To the surprise of the demonstrator, one point on the chart for averages was outside the control band. Here was a process said to be in a state of control but the test for control contradicted the advance billing. Naturally, students were eager to know whether the demonstrator switched bowls for the drawing which gave the maverick. If not, then "How often would at least one point be out?"

Quite precise answers to the above question and other related ones are now available and will be given.

Fundamentals

In order to clarify the exposition it seems useful to begin by reviewing a few statistical concepts and emphasizing certain distinctions. First the difference between state, hypothesis, and test is noted. An automobile battery may be in one of two states - dead or not dead. For some reason the hypothesis is put forward or the conjecture is made that the battery is dead. This hypothesis may be correct; one hopes it is wrong. Anyway, it is tested. Perhaps the test is to turn the ignition key to the proper position and to note whether or not there is a response. This test differs from the usual statistical test because there is no possibility of a Type I error for there will be no false indication of life if the battery really is dead. On the other hand there is the possibility of making a Type II error because, under certain conditions, one might have a live battery but get no response and so conclude that the battery was dead. In passing, it might be mentioned that, without a precise definition, there could be lack of agreement on the proper word to be used to describe the state of the battery.

A process, which has been set up to produce one ten-penny nail after another, may have been in a state of control during an eight-hour period or it may not. At each moment when a nail was to be produced the probability that the nail would be good might be the same. Furthermore, this marginal probability might be the same as the conditional probability when any sequence of good and bad nails had been produced previously. If both of the above conditions held, then the author would say that the process had been in strict control for the eight-hour period.

This description of strict control has been formalized by the following definition which was given by the author [10] at the Ninth Annual Convention of the American Society for Quality Control.

> "A process is in control in the strict sense during the time T_p, when
>
> (1) $f(x)$ is identical for every $x(t)$, with t in the interval, T_p, and
>
> (2) the set of random variables $[x(t)]$ are statistically independent."

It should be emphasized that for this first phase, control when no standard is given, $f(x)$ is not specified in advance. Furthermore, it is not known although, for a controlled process, it can be estimated.

Contrast the usual definition for control with respect to a given standard,

> "A process is in control at a time t, with respect to a standard, $f^s(x)$, if $f(x)$ is equal to $f^s(x)$."

Note that this definition refers to the state of the process at a specific time, t.

The test for control, when no standard is given, is well known. It is performed only after a set of samples have been drawn. It leads to the decision that the process is in a state of control only if all points are within their respective control bands.

On the other hand, a test for control with respect to a given standard is made after each sample is drawn. For example, twenty samples give twenty tests while, for the previous case, a set of twenty samples gives just one test.

Of course, as the author has suggested [10], it might be desirable to revise the procedure in the second phase and test for control with respect to a standard for a time period, T_s. Then the hypothesis of control will be accepted only if all points are within limits. Here, of course, the limits depend on the standard and can be set in advance. (It may be worth mentioning that a process may be in a state of control for a period, T_s, without being in a state of control with respect to the given standard.)

Usually, except, perhaps, for defect-per-unit charts, samples consist of measurements or counts for more than one unit. If the process is in a state of control, no logical difficulty is encountered. However, any hypothesized alternative to control should include the assumption that the process is in control during each minimum time interval during which a sample is taken.

THE CONTROL CHART FOR DEFECTS PER UNIT

Rather than following the chronological order of the investigative work, it seems best to begin with a discussion of power characteristics of the control chart for defects per unit. An industrial process is postulated such that, at any time t, the probability that a unit will have X defects is given by

$$(1) \qquad f(x) = \frac{(c')^x e^{-c'}}{x!} \qquad x = 0,1,2,\ldots$$

where c' may vary with time. It is assumed that k units will be inspected during the time T_p and that $c_1, c_2, \ldots, c_k$ will represent, in order of time, the numbers of defects found for the k units. Furthermore, $\bar{c}$ will symbolize the arithmetic mean of the c-values and UCL and LCL the upper and lower control limits.

While it is customary to use 3-sigma limits, there seems to be justification for using a-sigma limits in some cases, with a chosen to be some multiplier different from 3. Therefore, this procedure has been considered. Also the effect of using only a single limit has been investigated.

If, during the period T_p, a total of N units are produced, then the process will be said to be in a state of strict control if, and only if,

$$(2) \qquad f(x_1, x_2, \ldots x_N) = \prod_{i=1}^{N} f(x_i)$$

where the symbol on the left is the joint density for the indicated random variables and where $f(x_i)$ is given by (1) with x_i written for x and with c' the same for all x_i. ($\prod$, of course, means "product".)

The test for control is based on the use of the control limits to bound a control band. A set of k samples are taken in order of production and limits, given by the following formulas, are calculated.

$$(3) \qquad \begin{aligned} UCL &= \bar{c} + a\sqrt{\bar{c}} \\ LCL &= \bar{c} - a\sqrt{\bar{c}} \end{aligned}$$

If only an upper control limit is to be used, then the control band is bounded by x = UCL and x = 0. The latter line is part of the control band but the former line is not a part of it unless it coincides with the former line. If two control limits are to be used, then neither is a part of the control band except when both coincide with x = 0.

The hypothesis of control is rejected if one or more points fall outside the control band. Otherwise it is accepted. Then P_A, the probability of accepting the hypothesis that the process has been in a state of control during T_p, is given by

$$(4) \qquad P_A = \text{Prob}\,(LCL < x_1 < UCL, \ldots, LCL < x_k < UCL).$$

It should be emphasized that the test rejects the hypothesis if one or more points are outside the control band.

When the process is in a state of control the probability, α, of one or more points outside the band and so making the Type I error of rejecting the null hypothesis, is equal to $1 - P_A$. When the process is not in a state of control, the probability, $1 - \beta$, that the correct decision to reject the null hypothesis will be reached is $1 - P_A$. This is the power of the test. Obviously it will depend on the nature of the alternative state. (β, of course, is the probability of accepting the null hypothesis when it is false.)

In significance testing it has been the usual practice to fix α. Ordinarily the value chosen is 0.05. In hypothesis testing, where possible alternatives are considered in advance, it is customary to specify, also, the value of β for a particular alternative and then let the values of α and β determine the amount of information needed.

For the control chart, α is a function of c',k, and a. There may be an idea that α will be constant if k and a are fixed at some values such as 25 and 3. Such is not the case because α still varies with c'. In general, the computations necessary to get the exact value of α are lengthy. However, it is relatively easy to calculate upper and lower bounds for α which are close together. The procedure, which is based on the use of Bonferroni Inequalities is described in several reports [12,11,14,9]. As indicated in these documents, the upper bound used for α is close to the exact value of α when k is not too small. This upper bound, α_U tends to α as k tends to infinity. In any case, a study of the behavior of α_U is very helpful in revealing the performance of α.

Some indications of the variation of α_U with changes in c'k, and a, are given by Table I (compiled by extracting part of Table II from [14] p. 12 and rounding off the entries to 3 decimal places) and by Table II (obtained in similar fashion from Table IV-10 [9] p. 33). In Table I it might also be noted that all of the entries for k=25 exceed .05 and that the excess increases rapidly as c' decreases. For k=10, α_U has a maximum near $c' = 0.4$ and then decreases as c' increases. In every case α_U increases with k when c' is held constant. (Is this a desirable condition?)

Table II indicates the effect of changing the multiplier of sigma in the expression for the upper control limit. Study of the

TABLE I

UPPER BOUNDS FOR PROBABILITY OF TYPE I ERROR FOR UPPER THREE-SIGMA LIMIT CONTROL CHART FOR NUMBER OF DEFECTS, NO STANDARD GIVEN

c'	k=10	15	20	25
0.1	.036	.394	.360	.811
0.2	.086	.281	.315	.491
0.3	.109	.202	.301	.351
0.4	.111	.165	.262	.285
0.5	.104	.148	.228	.265
1.0	.070	.117	.164	.213
2.0	.039	.076	.106	.147
3.0	.032	.061	.087	.118
4.0	.030	.055	.080	.106
5.0	.026	.049	.072	.095
7.0	.024	.044	.063	.085
10.0	.021	.039	.055	.072
15.0	.018	.033	.049	.063
20.0	.016	.030	.044	.058

TABLE II

UPPER BOUNDS FOR PROBABILITY OF TYPE I ERROR FOR UPPER a-SIGMA LIMIT CONTROL CHART FOR NUMBER OF DEFECTS, NO STANDARD GIVEN, WHEN k=10.

	k=10			
c'	a=2.6	2.8	3.0	3.2
1	.117	.075	.070	.041
2	.093	.066	.039	.027
3	.082	.053	.032	.022
4	.072	.048	.030	.018
5	.070	.044	.026	.016
7	.064	.039	.024	.014
10	.059	.036	.021	.012
15	.054	.031	.018	.010
20	.050	.029	.016	.009

two Tables suggests that, if c' can be fairly well estimated in advance, it may be possible to choose (k,a) combinations for which the probability of a Type I error will be somewhere near .05. Some (c',k,a) combinations which approach this goal are given in [9], p. 36. Among the better ones listed are (1,5,2.5), (1,25,3.8), (2,15,3.2), (2,25,3.5), (4,20,3.2) and (4,25,3.3). It should be mentioned that only results for c_1=1,2,4 are tabled.

It might, perhaps, be suspected that the use of charts with both upper and lower control limits would have a considerable effect in altering α_{II}. Such is not the case unless c' is considerably larger than a $\sqrt{c'}$. It is interesting to know that, if a table like Table I were given here for charts with both upper and lower 3-sigma limits, only ten entries (all in the last three rows) would be increased and the largest increase would be .009 (at c'=20.0, k=25).

Turning now to the question of power, it is obvious that there a host of alternatives to strict control which might be listed. The first alternative to be considered here is one which has been called a single slippage. Suppose the total time interval has been divided into k sub-intervals and that the process is assumed to be in strict control during each of the sub-intervals. If, for k-1 of these sub-intervals, c' is equal to c'_1, but, for one of these sub-intervals, c' is equal to c'_2 which is different from c'_1, then the state of lack of control will be called a single slippage. Double slippages of equal amounts, a second alternative to be considered, is the case when c' is equal to c'_1 for k-2 of the sub-intervals and is equal to c'_2 (different from c'_1) for the other two. In each case the data for performing the test for control will consist of a sample of one unit from each sub-interval.

If the first alternative (single slippage) correctly describes the true state of lack of control then one of three mutually exclusive events can occur--

E_1: all points in the control band
E_2: the point corresponding to $c' = c'_2$ outside the control band (and, perhaps, other points also outside)
E_3: the point corresponding to $c' = c'_2$ <u>inside</u> the control band but one or more of the other points outside the band.

In the control chart operation either E_2 or E_3 is the signal to search for an assignable cause. The initial search, at least, probably will be concentrated on conditions characteristic of the sub-intervals generating the out-of-control points. In the case of E_2 the search will include the sub-intervals where the process has slipped and so the probability of E_2 is called "the probability of detecting a single slippage." The event, E_3, correctly signals

trouble but does not indicate the proper direction for the initial search. Therefore, the probability of E_3 often is referred to as "the probability of a Type III error." Obviously

$$(5) \qquad 1 - \beta = Pr(E_2) + Pr(E_3) = 1 - Pr(E_1) = 1 - P_A$$

Extensive tables for the probabilities for the detection of single and double slippages are given in the referenced reports. A few numbers have been chosen and rounded off in order to form the basis for Tables III and IV, which are given below. Study of Table III reveals that the power to detect a single upward slippage is somewhat greater for k=25 than it is for k=10. The improvement is not as great as might be expected in view of the fact that, for $c'=1$ and $c'=5$, α_U is between three times and four times as large for k=25 (see Table I). Furthermore, two and one-half times as many samples are required.

Inspection of Table IV reveals, as expected, that an increase in a is accompanied by a decrease in power. If both Tables III and IV are used in order to compare the power of the 2.6-sigma chart for k=10 with the 3-sigma chart for k=25, one may be surprised to find that the former has more power for the tabled combinations. This is the situation in spite of the fact that the Type I errors are less (see Tables I and II). The complexities of the general control problem are such that any firm generalization seem hazardous but results do suggest that it might be better to use samples from much fewer than twenty-five sub-intervals for the first phase. This is a matter, however, which requires further study. What does seem clear is the fact that, after the value of k is chosen, some attempt should be made to estimate c' in order that the multiplier of sigma can be so fixed that the value of α will be realistic.

A method of procedure has been obtained for calculating the power of the upper a-sigma limit chart to detect a double slippage of equal amounts. Detection here is to be interpreted as meaning that at least one of the slippages produces a point which falls outside the control band. Some calculations for cases where a=3, k=5, 10,20 have been made. For k=5, the power is less than the power to detect a single slippage for the (c'_1, c'_2) combinations which were used. For k=10 and k=20, the inequality was reversed. For k=10 and $c'_1=1$ the double slippage power values for $c'_2=1,3,5,10$ were .014, .227, .583, and .982, respectively. The corresponding single slippage power results were .007, .176, .507 and .950.

TABLE III

POWER OF THE UPPER THREE-SIGMA LIMIT CONTROL CHART FOR NUMBER OF DEFECTS, NO STANDARD GIVEN, TO DETECT A SINGLE SLIPPAGE.

	k=10		k=25	
c'_2	$c'_1=1$	$c'_1=5$	$c'_1=1$	$c'_1=5$
1	.007		.009	
2	.059		.078	
3	.176		.229	.000
4	.338	.000	.419	.001
5	.507	.003	.598	.004
7	.774	.027	.844	.039
10	.950	.185	.973	.244
15	.998	.676	.999	.756
20	1.000	.939	1.000	.965

TABLE IV

POWER OF THE UPPER a-SIGMA LIMIT CONTROL CHART FOR NUMBER OF DEFECTS, NO STANDARD GIVEN, TO DETECT A SINGLE SLIPPAGE WHEN k=10

	a=2.6		a=2.8		a=3.0		a=3.2	
$c_{\frac{1}{2}}$	$c_{\frac{1}{2}}$=1	$c_{\frac{1}{2}}$=5	$c_{\frac{1}{2}}$=1	$c_{\frac{1}{2}}$=5	$c_{\frac{1}{2}}$=1	$c_{\frac{1}{2}}$=5	$c_{\frac{1}{2}}$=1	$c_{\frac{1}{2}}$=5
1	.012		.007		.007		.004	
2	.085		.064		.059		.042	
3	.231	.000	.193	.000	.176		.140	
4	.412	.001	.365	.001	.338	.000	.288	.000
5	.586	.007	.541	.004	.507	.003	.454	.002
7	.831	.055	.803	.039	.774	.027	.736	.019
10	.968	.282	.961	.230	.950	.185	.938	.147
15	.999	.776	.998	.726	.998	.676	.997	.620
20	1.000	.967	1.000	.954	1.000	.939	1.000	.920

THE CONTROL CHART FOR FRACTION DEFECTIVE

Since an account of the author's research on power characteristics of the control chart for fraction defective has been issued by this Society as General Publication No. 4 [11], it seems unnecessary to devote much space to it in this summary. It might be mentioned, however, that some of the methods developed for it were useful in connection with the investigations described in the previous section of this summary. On the other hand, these later investigations led to the development of theory which can be helpful in extending the results previously reported. For example, a proof was obtained for the conjecture that, for $np_{\frac{1}{2}} = c_{\frac{1}{2}}$ and $np_{\frac{1}{2}} = c'_{\frac{1}{2}}$, the power of the upper three-sigma limit control chart for fraction defective, no standard given, to detect a single slippage converges to the power of the corresponding chart for defects per unit. Later this result was generalized so as to include lower limit charts and charts with both upper and lower limits.

THE CONTROL CHART FOR SAMPLE MEANS

As stated in the introduction, the need to investigate the power characteristics of the control chart in phase one was emphasized by the failure of the control chart for means to behave nicely in a public demonstration. Some time in the late forties, E. P. King, then a graduate student, became interested in the problem and, in 1951, completed a doctoral dissertation at Carnegie Institute of Technology, entitled, "The operating characteristic of the control chart for sample means when process standards are unspecified," Dr. King's original results and later extensions were reported in three published papers [5, 6, 7] and in a presentation to this Society at the Seventh Mid-West Quality Control Conference, at Indianapolis, in November, 1952. His work was especially noteworthy, not only because of the specific findings, but also because it initiated on a firm foundation a series investigations into the general problem now under discussion.

A very brief summary of King's dissertation can be written by paraphasing his own synopsis. He investigated the Type I and Type II errors for the control chart for averages with no standards given. The hypotheses assumed an underlying normal distribution but allowed random fluctuation in the process mean. Results were tabulated for cases ranging from two samples of two to twenty-five samples of ten. The effect of altering the traditional three-sigma limits was discussed and the probability of a Type II error was compared with the corresponding probability for the

traditional analysis of variance test.

King studied the control chart for means which is based on k samples of n units each and has a-sigma control limits. (King's notation has been changed slightly in order to avoid confusion.) He assumed that the ith sample of n, corresponding to the ith time interval, consisted of observations x_{ij} on random variables normally and independently distributed with mean μ_i and variance σ^2. Further, he assumed that the means were normally and independently distributed with mean zero and variance $\theta^2\sigma^2$. Explicitly

$x_{ij} = \mu_i + \epsilon_{ij}$ $(i = 1,2,\ldots,k_{ij} = 1,2,\ldots,n)$, with the μ_i NID$(0,\theta^2\sigma^2)$, the ϵ_{ij} NID$(0,\sigma^2)$ (all i,j), and the ϵ_{ij} and μ_i statistically independent.

It should be noted that (1) σ^2 is unknown, but constant, (2) there in no loss in generality in assuming that the expectation of μ_i is zero, and (3) this is the usual model for the one-factor analysis of variance with random components (Model II) with the exception of the assumption just discussed above.

For the control chart test based on sample means, King investigated the probabilities of the Type I and Type II errors for the null hypothesis, $\theta=0$, versus the alternative, $\theta > 0$. Writing

$$(6) \quad \beta(a|\theta) = \Pr(|\bar{x}_1 - \bar{\bar{x}}| < aA_2\bar{r}/3,\ldots,|\bar{x}_k - \bar{\bar{x}}| < aA_2\bar{r}/3)$$

and

$$(6') \quad \beta_0(a|\theta) = \Pr(|\bar{x}_1 - \bar{\bar{x}}| < a\sigma/\sqrt{n},\ldots, |\bar{x}_k - \bar{\bar{x}}| < a\sigma/\sqrt{n}),$$

where $\bar{r}$ is the mean of the k sample ranges, then

$$(7) \quad \alpha'(a) = 1 - \beta(a|0)$$

where $\alpha'(a)$ is the probability of the Type I error when a-sigma limits are used.

King found a way to calculate exact values of β_0 which proved feasible for k small but became so complex for k greater than 4 that, for the larger values, he contented himself with getting bounds for β_0. Also, by means of series expansions and numerical integration, he was able to get a series expansion for β interms of β_0 and central moments of the range distribution. For the mathematical details of the solution, the reader is referred to [5].

Values have been abstracted from King's dissertation in order to prepare the brief Table given below.

TABLE V

POWER, UNDER THE NORMALITY ASSUMPTION, OF THE THREE-SIGMA LIMIT CONTROL CHART FOR MEANS, NO STANDARD GIVEN, AGAINST MODEL II ALTERNATIVES

For k=25, only bounds are given

θ	k=4		k=25			
	n=5	n=10	n=5		n=10	
0	.99	1.00	.94	.97	.94	.97
.25			.81	.90	.64	.80
.50	.87	.74	.33	.56	.05	.23
.75			.03	.18	.00	.02
1.00	.51	.27	.00	.03		
2.00	.13	.05				
3.00	.05	.02				

Several comments regarding Table V and its implications might be made. In the first place, later work by Sorensen [15] seems to confirm King's conjecture that his lower bounds for β serve as good approximations to the exact values when k is large. Secondly, values for α can be obtained by subtracting from unity the β-values for $\theta = 0$. Thus α is .01 for k=4, n=5 and is between .03 and .06 for k=25, n=5. (Sorensen's work indicates that the exact value for the latter case is close to .07 so, perhaps, an unfortunate accumulation of rounding off errors occurred.) In the third place, just as for the two charts discussed previously, α increases with k. (But, for this chart, α does not depend on any unknown parameters.)

Impelled by C. C. Craig's paper [3], King compared the power of the test using the control chart for means with that of the analysis of variance test. Comparisons would not be valid, of course, unless the α for the F-test were set at the same level as for the control chart test, and, as has been noted, the value is not, in general, .01 or .05. Using the Incomplete Beta Function and the work of Ferris, Grubbs and Weaver, proper adjustments were made and comparisons were obtained. From his results, King came to the expected conclusion that the analysis of variance is definitely more powerful under this particular alternative hypothesis. He noted, however, that the difference in power appeared to decrease markedly with increasing θ. More extensive comparisons are given in King's dissertation. It may be sufficient here to note that, for $\theta = 0, .25, .50, .75$, and 1.00, the power values for the control chart test, (as calculated using the lower bound for β), for the case k=25, n=5, are .06, .19, .67, .97, and 1.00 respectively, while the corresponding values for the analysis of variance test are .06, .28, .85, .99, and 1.00.

Recently, Sorensen [15] decided to attempt a verification of a conjecture of Wallis and Roberts, [19], pp. 498-501, and others that, for some alternatives the test using the control chart for sample means, no standard given, is more powerful than the analysis of variance test. He postulated the same mathematical frame work and sampling procedure as King did but replaced the alternative hypothesis that the sub-interval means were $NID(0,\theta^2\ \sigma^2)$ with $\theta > 0$, by alternatives which consisted of various types of slippages. His five alternatives were:

1. A single mean has slipped,
2. Two means have slipped equally to opposite sides,
3. Two means have slipped equally to the same side,
4. The means follow a linear trend,
5. Half the means are at each of two levels, (If k is odd, one mean

is assumed to be midway between the two levels)

Sorensen attacked again the general problem of getting exact values for α and β. The equations

(8) $\beta(a \mid \mu_1/\sigma, \ldots, \mu_k/\sigma) = \Pr(|\bar{x}_1 - \bar{\bar{x}}| < aA_2\bar{r}/3, \ldots, |\bar{x}_k - \bar{\bar{x}}| < aA_2\bar{r}/3)$

and

(8') $\beta_0(a \mid \mu_1/\sigma, \ldots, \mu_k/\sigma) = \Pr(|\bar{x}_1 - \bar{\bar{x}}| < a\sigma/\sqrt{n}, \ldots, |\bar{x}_k - \bar{\bar{x}}| < a\sigma/\sqrt{n})$.

An approach similar to King's was used in obtaining an expression for β which depended on β_0. For $k > 3$ difficulty was encountered in getting β_0 for the alternatives considered. The problem was solved by utilizing a series expansion for which convergence was established with considerable difficulty. It is expected that the details of the mathematical analysis will be published elsewhere.

A state of strict control exists when $\mu_1 = \mu_2 = \ldots = \mu_k$. The alternative of a single slippage is specified when μ_1 is equal to $\mu + \delta\sigma$ and the other $k - 1$ means are equal to μ. Obviously, the direction of the slippage depends on the sign of δ. As before, there is probability that the mean of the sample from the ith sub-interval will be outside the control band and this is designated as the probability of detecting the single slippage. But the power is the probability that, for this alternative state, at least one sample mean will be outside the control band. Therefore, for k more than 2, the latter probability will exceed the former. For the three-sigma limit chart, both of the above-named probabilities have been calculated for a useful variety of combinations of k and n. It is hoped that the extensive tables and graphs of the results will be published in Industrial Quality Control. For the present purpose, a few numbers have been abstracted in order to form the Table given below. These numbers give the correct indications of variation (except for rounding off effects). Note that (1) α increases with k but decreases with n, (2) power (a) is greater than power (b) and both increase with k, n, or δ.

As noted above, α varies with k and n. It also decreases with a. For a set of (k,n) combinations, values of a for which α is equal to .05 have been calculated. For (10,5), (20,5), (25,5), the values are 2.79, 3.03, and 3.10 respectively. For n large (i.e. sigma known), the values k=10, 20, 25 are 2.65, 2.94, 3.02. For each such a-value and (k,n) combination it is possible to calculate the power to detect each of the five kinds of slippages listed above. Therefore, power comparisons with the analysis of variance test can be made.

TABLE VI

POWER OF THE THREE-SIGMA LIMIT CONTROL CHART TEST FOR SAMPLE MEANS, NO STANDARD GIVEN, TO DETECT (a) LACK OF CONTROL, (b) THE SINGLE SLIPPAGE, WHEN THERE IS A SINGLE SLIPPAGE OF SIZE $\delta\sigma$

	k=10				k=25			
	n=5		n=10		n=5		n=10	
δ	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
0.0	.03		.02		.07		.06	
0.5	.05	.02	.07	.05	.09	.03	.12	.07
1.0	.19	.16	.45	.44	.25	.20	.54	.51
1.5	.52	.51	990	.61	.61	.59	.94	.94
2.0	.85	.85	1.00	1.00	.91	.90	1.00	1.00
2.5	.98	.98			.99	.99		
3.0	1.00	1.00			1.00	1.00		

Tang [17] found that the operating characteristic for the Model I alternatives can be expressed as a function of the non-centrality parameter

$$(9) \quad \varphi^2 = (n/k\,\sigma^2) \sum_{1}^{k} (\mu_i - \bar{\mu})^2$$

Ura [18] has published tables from which values of the above parameter, corresponding to $\alpha = .05$ and $1 - \beta$ .90, can be obtained. Having obtained φ^2 for a value of n and k, the values of δ for each of the five types of slippages can be calculated. Thus the desired comparisons can be made. For n large and k equal to 10, the power to detect lack of control for the five types are .97, .92, .88, .63, and .54. For k=25 the corresponding values are .99, .95, .93, .53, and .44. These values, of course, should be compared with .90, the value for the analysis of variance test.

THE CONTROL CHART FOR RANGES

In connection with a technical report [12] and in the preparation of his doctoral dissertation, R. W. Kennard investigated the properties of control charts for dispersion. This work, of course, included study of the control chart for ranges, no standard given. After careful analysis, he concluded that, in order to make any practical conclusions about a control chart using ranges, it seemed necessary to use experimental sampling. At that time, the amount of time available for experimental sampling was not great and the conclusions must be viewed as very tentative. From his work it seemed that α is small (practically zero for 3-sigma limits unless k exceeds 15) but that it increases with k, that the test has better power against the alternative of a single upward slippage than against general heterogeneity, and that this power increases with k. Also he thought that the test might not be robust and that non-normality rather than lack of control might sometimes be the cause of an indication of lack of control.

FINAL REMARKS

In the transition from phase one to phase two of the control operation, it is necessary to stop and establish standard values for the parameters. If the process is in a state of control and the test for control leads to the correct decision then parameters are estimated. If, on the basis of these estimates, the process is deemed to be satisfactory then it is a usual procedure to use these estimates as standard values. In [12], pp. 68-71, it was pointed out, if the process was in a state of control when the k samples were taken and if it remains in control when the $(k + 1)$st sample is taken, (i.e. the $k + 1$ samples are from the same normal universe), then the probability that the $(k + 1)$st sample falls within the control limits for sample means can be calculated by use of the t-distribution with $k(n - 1)$ degrees of freedom, if the control limits were based on a variance estimate obtained by pooling within-sample variances. If limits were based on mean range, then, as noted, the tables of Lord [8] can be used. Conversely, of course, the original limits can be set so that the probability is fixed at a desired level.

The above statements might serve as a veiled warning against setting standards on the basis of too few data. Such a warning may be necessary in view of the emphasis in some of the previous sections on the fact that an α of a desired size can be gotten from various (k, n) combinations by the proper choice of a, the sigma multiplier. As usual there is an upper limit on the value of a set of data. It is useful, however, to have some idea what that limit might be.

REFERENCES

1. ASTM Manual on Quality Control of Materials, American Society for Testing Materials, Philadelphia, 1951.

2. Control Chart Method of Controlling Quality During Production, American Standards Association, 1942.

3. Craig, C. C., "Control charts versus analysis of variance in process control by variables", Industrial Quality Control, Vol. 3, No. 4, (Jan. 1947), pp. 14-16.

4. Ferris, C. D., Grubbs, F. E., and Weaver, C. L., "Operating characteristics for the common tests of significance", Annals of Mathematical Statistics, Vol. 17 (1946), pp. 178-197.

5. King, E. P., "The operating characteristic of the control chart for sample means", Annals of Mathematical Statistics, Vol. 23 (1952), pp. 384-395.

6. King, E. P., "The operating characteristic of the average chart", Industrial Quality Control, Vol. 9, No. 3 (Nov. 1952), pp. 30-32.

7. King, E. P., "Probability limits for the average chart when process standards are unspecified", Industrial Quality Control, Vol. 10, No. 6 (May 1954), pp. 62-64.

8. Lord, E., "The use of range instead of standard deviation in the t-test", Biometrika, Vol. 34 (1947), pp. 41-67.

9. McCue, E. B., and Olds, E. G., "A Further Investigation of the Power Characteristics of the Control Chart for Number of Defects, No Standard Given, Technical Report No. 5, Office of Ordnance Research, U. S. Army, Cont. DA-36-061-ORD 447, Carnegie Institute of Technology, Pittsburgh, May 1959.

10. Olds, E. G., "The nature of the standard control chart, and some of its competitors", Industrial Quality Control, Vol. 13, No. 4 (Oct. 1956), pp. 4-8.

11. Olds, E. G., The Power to Detect a Single Slippage and the Probability of a Type I Error for the Upper Three-Sigma Limit Control Chart for Fraction Defective, No Standard Given, ASQC General Publications, No. 4, American Society for Quality Control, Milwaukee, April 1958.

12. Olds, E. G., and Kennard, R. W., On the Determination of Significance Levels of Control Charts for Continuous Variates, Technical Report No. 2, Office of Ordnance Research, U. S. Army, Cont. DA-36-061-ORD 447, Carnegie Institute of Technology, Pittsburgh, June 1955.

13. Olds, E. G., and Knowler, L. A., "Teaching statistical quality control for town and gown", Journal of the American Statistical Association, Vol. 44 (1949), pp. 213-230.

14. Olds, E. G., and McCue, E. B., The Power to Detect a Single Slippage and the Probability of a Type I Error for the Upper Three-Sigma Limit Control Chart for Number of Defects, No Standard Given, Technical Report No. 4, Office of Ordnance Research, U. S. Army, Cont. DA-36-061-ORD 447, Carnegie Institute of Technology, Pittsburgh, June 1958.

15. Olds, E. G., and Sorensen, F. A., Operating Characteristics of the Control Chart for Sample Means as a Test for Model I Alternatives, Technical Report No. 6, Office of Ordnance Research, U. S. Army Cont. DA-36-061-ORD 447, Carnegie Institute of Technology, Pittsburgh, May 1959.

16. Scheffé, H., "Operating characteristics of average and range charts", Industrial Quality Control, Vol. 5, No. 6, (May 1949) pp. 13-18.

17. Tang, P. C., "The power function of the analysis of variance tests with tables and illustrations of their use", Statistical Research Memoirs, Vol. 2 (1938) pp. 126-149.

18. Ura, S., "A table of the power function of the analysis of variance test", Reports of Statistical Applications Research, Union of Japanese Scientists and Engineers, Vol. 3, No. 2 (March 1954), pp. 1-6.

19. Wallis, W. A., and Roberts, H. V., Statistics: A New Approach, The Free Press Glencoe (1956), pp. 498-501.

20. Working, H., "Statistical quality control in war production", Journal of the American Statistical Association, Vol. 40 (1945), pp. 425-447.

LIFE TEST - SOME PRACTICAL CONSIDERATIONS

H J Davis, Raytheon Company
B P Goldsmith, Northeastern University

ABSTRACT

The design of any worth-while specification must satisfy clearly defined and realistic objectives. The authors review current needs and specifications in the area of life testing components in the electronics industry, and propose a practical technique to satisfy changing requirements. The proposal is based on actual performance and has been shaped to make use of readily available techniques and equipment.

Introduction

Life testing is so universally required by today's specifications, particularly on components for military use, that it is commonly accepted as a way of life by the manufacturer and his engineers. With the continually increasing emphasis on reliability of components and systems, reliability and quality assurance engineers are advocating new life tests, or extension of existing tests, with the fervent hope of attaining higher orders of reliability as a consequence of these new or more severe tests. Practical consideration of the aims and means of performing tests at first may be overlooked.

It is the intent of this paper to review the objectives of life tests in the hope of stimulating an informed and realistic review of current proposals.

Objectives

In view of the current stress on life testing, which gives every indication of being intensified in the future, it will be well to review briefly the objectives and general means of performing life tests in order to appraise more intelligently the various schemes now in use or being proposed. The basic aims of life testing and consequent costs are all too often neglected in the flurry of running tests and accumulating data.

The first step in designing or selecting a life test plan is to review and clearly define the objective(s) of the test. While this would seem obvious, discussion of such items as sample sizes, failure rates, and methods of measurement have often become so spirited that the basic objective of the test is no longer dominant.

ASQC LCS Code: 815, 840:60:436

Life tests may be generally classified as having one or more of the following purposes:

1) Exploration of product characteristics
2) Lot-by-lot product acceptance
3) Qualification of a design and/or a process
4) Qualification of a manufacturer (skills and equipment)
5) Surveillance of (3) and (4) above on a continuing basis

Life tests may be required in order to complete the description of either an existing or a new product and so permit full engineering evaluation. Exploratory tests may seek to determine information about electrical, mechanical, or thermal characteristics of a component.

On the other hand, for an established product it is frequently necessary to determine whether a particular, discrete, lot meets specified standards of reliability. Further, there may be a need to evaluate designs of a component, or the processes by which the components are produced, in order to see whether the designs and/or the processes are adequate for producing a reliable product. Similarly, when a design or a process has been established as being capable of producing a reliable component, it is important to know whether a particular manufacturer is capable of effectively making the product in accordance with approved design and process specifications. Finally, we may wish to know how a particular design or process performs over an extended period of time, or how stable the product from a particular manufacturer is in a long sequence of lots.

While a life test may have only a single objective, it is evident that in many situations it would be desirable to run tests for more than one reason. Accordingly, a given life test plan may be designed so as to embrace two or more of the general objectives listed earlier.

Evaluation of Performance

After the basic objective of the test has been resolved and clearly defined, it is necessary to establish a method of evaluating the component or equipment under test. Some recognized methods are:

Functional test under in-use conditions
Simulated functional test under in-use conditions
Accelerated functional test under in-use conditions
Environmental profiles

Since actual performance in field equipment is the final criterion of performance of any component, life tests under actual use conditions would seem to be the most desirable method of evaluation. Yet the wide variety of conditions to which identical types of equipments or systems may be subjected and the variations introduced by operating personnel make tests most difficult to handle. The problem of accurate data reporting is in itself a most formidable obstacle which has received extensive attention, but which is by no means solved. A recent ARINC report (1) states: "50 to 70 % of tubes removed were unnecessary removals in the sense that the equipment would have continued to operate satisfactorily if they had not been removed."

Manufacturers of components, and in particular those of electron tubes, are painfully aware of the fallacies which may be introduced by taking field failure reports at face value, since often more than half of the components returned as "defective" are indeed found to comply with all specification requirements. This is usually due either to failure in equipment as a result of change in characteristic(s) which are still within component specifications, or to poor repeatability of some types of characteristic tests. Therefore field life tests must be handled with great caution, although this aspect of testing should be carefully considered in conjunction with other types of tests.

Simulated Functional Test

Since the hazards of loosely coordinated field testing are so great, simulated functional tests are often used. A far greater degree of control can be exercised on such tests, since they are usually much more compact, and subject to close supervision of the testing agency. One problem here is to ensure that the simulated conditions do in fact represent field conditions, so that conclusions drawn from simulated tests may reasonably be extended to predict field performance. It is frequently impossible to extend to specialized and military applications the large-scale chassis testing which has become so common in the TV and radio components industry.

Accelerated Functional Tests

Because of the long duration of tests necessary on components with low failure rates, attempts are often made to accelerate the tests by increasing the severity of electrical, mechanical, or thermal test conditions. Here the hoped-for objective is to get failure information which can be correlated with field performance under more normal conditions. Such techniques have been successfully applied to the testing of resistors and capacitors, but have been inadequate for many other components and assemblies, including electron tubes.

Environmental Profiles

Specifications on many components now call for tests under a variety of conditions, where the environment is changed according to a predetermined pattern. This pattern usually involves programmed changes in physical environment and electrical conditions, over the domain of parameters to permit prediction of performance in a broad range of field conditions. This is particularly true of military and airborne items, which must be capable of withstanding extremes of temperature, pressure, and other environmental stresses not usually encountered by commercial equipments.

Environmental profile tests may be considered forms of simulated or accelerated tests and require all the caution in design and interpretation previously discussed in this connection.

Selection of Critical Parameters

The choice and unequivocal definition of the parameters to serve as the criteria of performance is most important. A clear definition of a parameter almost always requires that it contain a clear description

of a valid test of that parameter. Unwieldy and unenforceable specifications have resulted from hasty attempts to catalogue all parameters suspected of being of interest in one or another of the long list of applications of many components. Fortunately, in many cases a careful choice of only a few parameters will serve as a reliable index of performance.

Pareto's established rule in economics has long been found to have a direct counterpart of wide applicability in engineering and industry. To paraphrase Juran's (2) statement of Pareto's rule: "In any series of failure causes, a small number of identifiable causes always accounts for a large fraction of the failures"!

Thus, measurement of plate current, or of transconductance, or of the change in these parameters, will quite adequately describe the performance of electron tubes in most applications. In general, success depends on the selection of the fewest characteristics which will measure the behavior of the component.

In any case, there is little justification in adding end-points and tests to a specification unless such additions contribute a significant improvement to the test's total discrimination.

Choice of Time Interval

Duration of testing presents another nice problem in balance. If the test is conducted for too brief a time, it may not generate sufficient information, or may invite unwarranted extrapolation about performance over a longer period. This is especially true if the mode of failure changes during life. On the other hand, if the test is unnecessarily protracted, the delay in taking proper action, or in releasing the product, may be extremely costly. The choice of time interval between observations is a further consideration. If the mode of failure is important, the intervals must be sufficiently short to enable a valid estimate of the failure distribution.

Cost Factors

Considerations affecting costs of life testing should certainly include:

Dollar costs of equipment and performance of tests
Intangible costs in reputation, morale, engineering leadership, etc.
Cost of information (or lack of it)
Excessive complexity, lack of flexibility and adaptibility
Lost time.

Life tests may be extremely expensive beyond the direct dollar costs of testing. Such costs may not be anticipated by manufacturers who have been stampeded into seeking reliability beyond the current state of the art. It is usually absurd to require for commercial applications the expenses that are justified only for equipment destined for use in outer space or in submarine cables.

An important consideration in the choice of life test specifications is the availability of test equipment to gather the required data. Test equipment costs are not confined to, but do include design, rental, acquisition of life test racks, test and read-out equipment, wages for operating and maintenance personnel, acquisition and operation of ancillary

equipment (air conditioning, special monitoring equipment, etc.). Design and development costs may be especially important to environmental profiles and ultra-long life specifications. Such specifications may require not only the best components and equipment commercially available but incur additional costs in making a real advance beyond the current state of the art. The cost of developing specialized measuring and testing gear may exceed the cost of improving the product, unless improvements have wider applicability than the product under current investigation.

Intangibles include the reputation of the producer. If test facilities are inadequate, of if there is a record of failure to meet a life requirement, this may reflect unfavorably on the entire product line. Similarly, lack of sufficient information may prevent a producer from substantiating claims for a reliable product. Unanticipated high costs of an extensive test program may preclude a profit even on a product of demonstrated high reliability. Poorly chosen test intervals or duration, previously mentioned, may adversely affect all the other cost factors.

The life test requirements, the available equipment, and operating personnel are inextricably interrelated. It is unrealistic to consider one without reference to the others. It is crucial to consider the cost and gain in information as well as in dollars in the listing of end points and characteristics to be monitored. Characteristics should not be added lightly to a specification in the mere hope that sooner or later mountains of data may suddenly turn to pay gold.

The choice of characteristics and time intervals have the greatest influence on complexity of administration and equipment.

A high degree of complexity inhibits flexibility in expanding a successful specification to the rest of the product line. Large sample sizes in general will make more automatic testing and equipment desirable, while an increase in a number of end points will be a deterrent to such progress. It is unlikely that any amount of automation or integration of test equipment with electronic data processing equipment will soon reduce the necessity for thoughtful simplification of procedures and specifications: if anything, current equipment and practice requires that automation be preceded by a greater degree of simplification and greater sophistication in design than most manual systems would have required.

Dynamics of Test Equipment

There is no static life test - all life tests are dynamic with respect to test conditions and environment, as well as time.

Some variations in test conditions and environment are carefully programmed into the testing routine, e.g. heater cycling, vibration and fatigue tests, temperature and humidity cycling, and thermal strain tests. Unintentional variations, however, frequently may be as large and as decisive to the performance of the component or equipment as the intentional variations. Poorly regulated electrical supplies may have a disastrous result on long-life performance, equipment may be unwittingly exposed to thermal shock (open window near temperature sensitive equipment), and rough handling may be more traumatic than either shock test or use environment. ARINC• and similar investigations have shown the vital importance of correct installation of components under test, as well as careful maintenance and use of test

equipment. A life test schedule requiring frequent readings produces in some components mechanical and glass failures which never would have occured under even the most severe in-use conditions.

A marked increase in sample sizes tends to down-grade the care and attention devoted to life test sockets and terminals. Sockets and connectors which were designed to be used for only a small number of make-and-break connections are subjected to hundreds, or even thousands, of these cycles during life test. Reliable data requires careful routine calibration and maintenance of test and measurement equipment for life tests no less than for any other engineering activity.

Sampling

This subject has received much attention from theoreticians, statisticians, and engineers. Unfortunately, much of the attention lavished on life test sampling is the result of frustration in attempts to state concise objectives of tests and inability to fix accurate costs. The consequence has been the design of sampling plans for general application without benefit of specific supporting data on the state of the art and progress attainable at a given cost. Such flight to theory unencumbered by fact, results in computational procedures intended to squeeze the last dram of information from life test data without regard to testing errors and natural variability inherent in present product and test equipment.

In applying any life test plan operating personnel must have a clear understanding where and how to draw the sample, when and how to test it, and how to report results. It is not enough to say "the sample shall be drawn in a random matter so as to represent the lot whose designation it bears." A sample drawn from a particular sequence of product at the end of assembly will not yield the same results as the same sample drawn from the same sequence of product after testing, or after aging or after packing, if for no other reason than that after each step the product is somewhat older. If the product has noticeable infant mortality, small delays in testing may have a marked effect on results.

Interpretation of Life Test Results

The purpose of life tests is to furnish information for engineering judgment of the product or process under test. The objectives of such evaluation have been outlined earlier in this paper. It is important to keep in mind that the statistical techniques and statistical considerations in design, sampling, and analysis are merely means to achieve the desired end, namely engineering evaluation. Under no circumstances should the end be subordinated to the means. This has been a temptation in other fields, as in indiscriminate use of control charts in production. Life tests in themselves, as control charts in themselves, are sterile and unproductive without mature evaluation of the intelligence they contain.

Information, interpretation, and evaluation are all merely steps towards feedback and action. A smooth flow can be facilitated by a realistic life test specification.[1]

Feedback

Life test is a time-consuming, slow process; therefore, it is all the more important to make fool-proof arrangements for clear and concise

transmission of significant results. Life test results are of immediate importance to factory engineering, quality control, production control, and production departments, significant changes in level or persistent trends are of concern to and may require participation by the sales department, Research and Development groups, and perhaps some officers in senior management.

The results of life tests and the implications of new specifications will increasingly involve members of the company who have never before been concerned with such esoteric activities. It is important that the initiated make a continuing serious effort to inform their colleagues and to keep them abreast of developments, without resort to unnecessary technical jargon or detail.

Representative current specifications

Routine life tests on production runs have become common only since World War II. Careful design, collection, and analysis of such data has come of age even more recently. Early specifications were hopefully intended to indicate the deterioration rate of components or equipment under test and to demonstrate that they would perform in service for "some reasonable time". For example, the pioneer JAN life test specifications required only that a 5-tube sample earn 80% of the credit hours attainable in 500 hours of testing under fairly conservative conditions. The low precision of such small sample predictors of lot performance was well matched by the limited ability to describe end-use conditions. This latter situation gave rise to the apt monicker "the universal unreliable component."

There was little point in drastic increases in expenditures for samples or testing until application requirements could be more clearly defined, in respect to conditions, parameters, and duration. The first step in this was the identification and separation of product intended for entertainment end-use from that destined for airborne, military, and certain specialized industrial applications in which the dire consequences of failure justified a higher product cost. The remainder of this paper is concerned with this latter category.

The early MIL specifications are representative of early efforts to write lot-by-lot product acceptance specifications to suit high-performance applications. This restriction made possible the design of a number of environmental tests (e.g. high temperature, humidity, and altitude tests). The duration of tests also grew: tubes were now left on life test for 1000 hours as a matter of routine.

These specifications were almost exclusively intended for lot-by-lot product acceptance, with only incidental surveillance of processes and manufacturers.

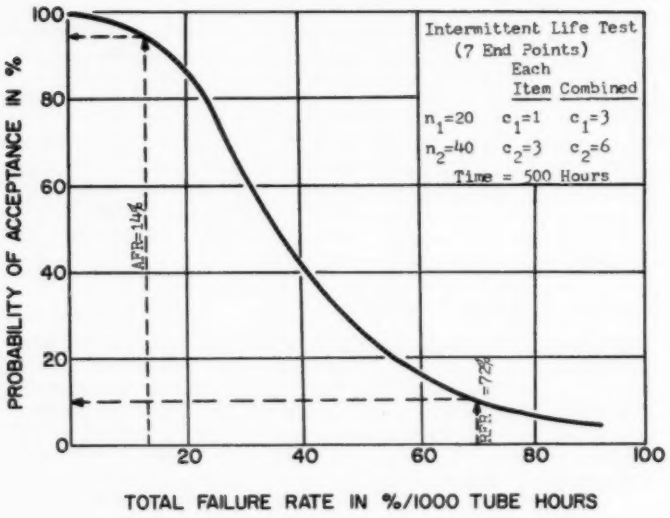


FIGURE 1 -- Approximate Operating Characteristic for a MIL-E-1 Intermittent Life Test with 7 End Points

A recent approach to obtaining better reliability data on electron tubes has been advanced by an ad hoc committee under the auspices of The Office of the Assistant Secretary of Defense for Supply and Logistics (OASD) 3,. This committee considered several plateaus of reliability, Level I being the level attainable in the current state of the art. Subsequent le vels were to carry well beyond capabilities attainable in the immediate future.

The OASD proposal for tubes contemplates a new high-reliability life test sample for Level I of n-80, with acceptance number c-2, and a test duration of 1,000 hours. The number of end points has been reduced to facilitate handling of the comparatively large sample size, and to enable direct reading without removing the tubes from the rack. (see Figure 2).

The first marked deviation from lot-by-lot product acceptance is contained in the WADC Capacitor specification of JAN 1958 (MIL - C - 262443).

This document has several objectives:

Lot acceptance
Qualification of the design and process
Continuing surveillance of a manufacturer
Encouragement of the manufacturer to continue to reduce the failure rate of the product.

The sampling rate, as well as the sample size required for each production lot, depends on the total quantity contained in all production lots

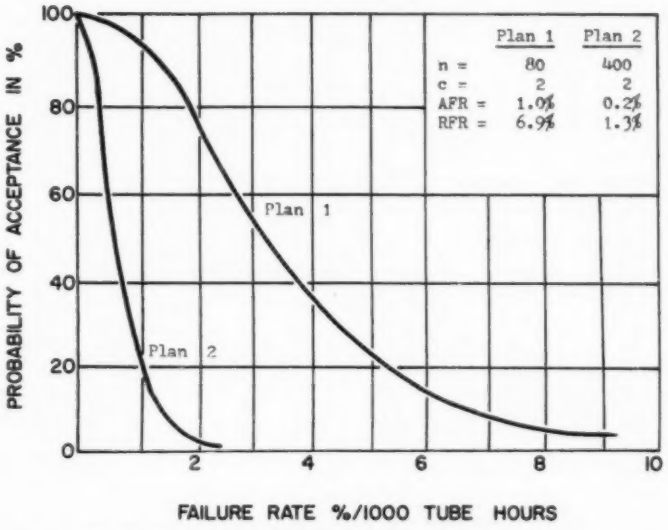


FIGURE 2 -- OASD High Reliability Life Test Operating Characteristic Curves

submitted during the preceding 30 production days. For very small production quantities the sampling rate is 30% of any one production lot; this sampling rate drops gradually with increasing production quantities, leveling off at 5% of any production lot from a large run.

This specification provides for the rejection of individual lots whose life test sample have excessive failures, and renegotiation of the contract if there is evidence of long term deviation from the nominal acceptable failure rate. (see Figure 3)

It is evident that evaluation of a continuing series of life tests, as recorded on the production log and the cumulative plot, affords a means of rating the manufacturer, and performing continuing surveillance on his product and facility.

Multiple objectives require a flexible plan

The demand for higher reliability continues to intensify; it is becoming increasingly apparant that no single set of checks is adequate to give sufficient quality assurance: the design, the processes, the subassemblies, the manufacturer, and the end product are all susceptible to evaluation, and modern needs compel us to base conclusions on the sum of all available information.

An important part of the proposal is a continuing log of test results and a plot of the total number of failures in the last four samples. The following plan was developed to meet the needs for high reliability tubes of WADC, and to consistent with current desires of DOD (Table I, Figure 4).

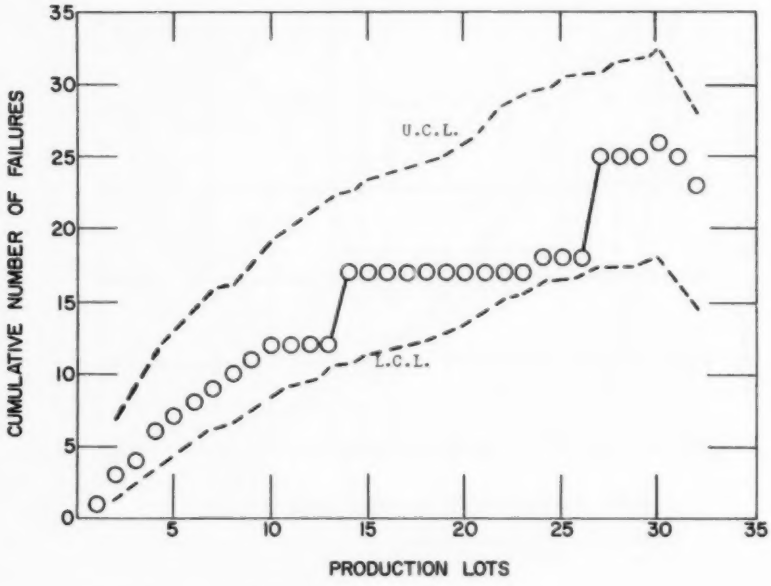


FIGURE 3 -- Typical Chart of Cumulative Defectives (MIL-C-26244 Capacitor Specification)

Sample sizes depend only on the degree to which current life test data is consistent with the specified Acceptable Failure Rate (AFR). If the observed rate is significantly higher than the AFR, the sample size increases to a present maximum; failures in excess of the limit on failures in a single sample result in lot rejection. Continued results in the neighborhood of the AFR permits successive reductions in sample size until a preset minimum is achieved. An average sample failure rate significantly below the AFR is a qualification for renegotiation of the contract.

This plan differs from both the OASD and WADC proposals in that the sampling rate is entirely independent of production quantities, and puts a constant bound on the confidence level associated with the current estimate of the failure rate.

This plan has been designed to achieve the following objectives:

Generation of reliability data in a routine and orderly manner, subjected to regular engineering review.
Evaluation of a particular manufacturer of a given design.
Surveillance of product, as well as lot-by-lot acceptance.
Constant incentive for the manufacturer to strive for quality improvement.
Adaptability to meet changing reliability requirements without major redesign of the plan.

TABLE I -- Log of Life Test Results on a Typical Miniature Tube Type (New Proposal)

LOT	SAMPLE SIZE	CUMULATIVE SAMPLE*	FAILURES	CUMULATIVE FAILURES*	UCL	LCL
1	100	100	1	1	-	-
2	100	200	1	2	5.5	-
3	60	260	1	3	6.5	-
4	60	320	0	3	7.5	0.5
5	40	260	0	2	6.5	-
6	40	240	2	3	6.5	-
7	40	220	0	3	5.5	-

* In Last 4 Samples

Confidence Limits = $M/N \pm 1.65\sqrt{M/N}$

where, M = Total Number in Last 4 Samples, and Allowable

Failures = $1/N$ (N = 75 in this example)

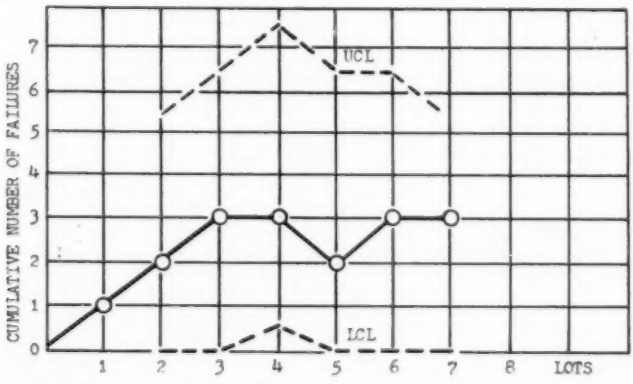


FIGURE 4 -- Chart of Cumulative Failures (Last 4 Lots) on a Miniature Tube Life Test Series

This plan may be easily adjusted to meet the demands of a variety of situations. For example, if history over a number of lots demonstrates that the product can easily meet the initially established failure rate, it is simple matter to renegotiate the AFR and/or price. Again, to increase the precision of samples or to accumulate data more rapidly, the sample sizes may be increased or additional larger samples may be tested. It it is desired to increase the pressure for reliability improvement, the confidence level and/or sample size may be adjusted. When a cumulative sample over a larger segment of product is needed, it is a simple procedure to increase the number of consecutive lots over which the sampled data is summed.

This basic framework for life testing and reporting results has the flexibility to meet a wide range of requirements, and brings with it the advantages of standardization: consistency in performance and administration, simplified training, evaluation, and comparison.

The proposal is not limited to any arbitrarily chosen failure distribution. Failure rates and testing schedules may be based on the Normal Weibull, or exponential distribution. without making any changes in routine instructions.

References

(1) Aeronautical Radio, Incorporated, Publication # 101-16 -144, September 15, 1959.

(2) Universals of Management Planning - J M Juran, American Management Association. The Management Review, November, 1954.

(3) Department of Defense, Offices of the Assistant Secretaries of Defense (R&E, S&L) Ad Hoc Study Groups on Electronic Parts Specifications Management for Reliability (1958-1960).

GOOD ENOUGH - ISN'T GOOD ENOUGH

Bernard I. Raysor, Works Manager
Harrison Radiator Division, Lockport, New York

"Good Enough - Isn't Good Enough." We all used to think that if we had attained 99% of anything we were outstanding. We have become accustomed to thinking in terms of Quality Control. Reliability was accepted as just a word. Now, the Reliability concept has opened a new field of vision - a field of predictability - a true measure of performance.

When new and old products are fighting for every sale possible mistakes can be more than damaging, they can be fatal. Not only are problems more numerous and more complex today than ever before, they are more critical.

It has been pointed out that with a parts reliability of 99.44% not a single Atlas missile would hit its mark. Certainly this points up what we may have thought was "good enough - isn't good enough." It has been further pointed out that to have 9 out of 10 Atlas missiles hit their mark, their parts reliability must be 99.99996%. This requirement is frightening when you realize that our very survival depends upon this kind of perfection. It is true that this kind of reliability started with the missile program, but has quickly spread to all of our Industry.

Our General Manager, Lawrence A. Zwicker, has said - "Reliability is not just a word. It is a goal that must be reached. It must influence our thoughts and govern our actions during every working hour."

We at Harrison recognize this and our Management is putting forth the effort to improve our product quality with reliability evaluations, research, product design, tool design, manufacturing processes, and last but not least, quality control and people.

It is these last two factors that I would like to discuss today. We are convinced that you can have perfection in all other factors, but with a lack of what we at Harrison call "the quality attitude - good enough isn't good enough," you will not have product reliability.

All of us have, I am sure, at one time or another defended our goofs with "it's good enough - it was only a small error - it is less than 1% - man, what do they expect - to err is human - to forgive is divine." This attitude seems reasonable to us in light of the fact that all through our youth in school we were permitted to perform at 75% or even less. And all of a sudden we were violently awakened to the fact that in Industry we were expected to produce 100% of the rates established for the operation, 100% quality workmanship, 100% safety, 100% this, 100% that. With our early training and attitude, is it any wonder that as a customer occasionally we find someone has goofed on the products we buy? Again, "good enough - isn't good enough."

Certainly, all of us recognize that quality control, in itself, controls nothing but is a means of communicating and pointing up conditions, good or bad, in a factual manner. In all instances, the information provided requires human action. An axiom that I think points this up very well is -

To look is one thing
To see what you look at is another
To understand what you see is a third
To learn from what you understand is still something else
But to act on what you learn is all that really matters.

Can we afford to ignore this axiom?

The warranty expense being paid by Industry today will answer this for you very quickly. Wouldn't a dime or a dollar spent in warranty repair have been much better spent in preventive action prior to shipment to the customer? The money spent in the plant has not cost the manufacturer his good quality reputation, which is almost unsalvageable and for which there can be other serious monetary losses, as a matter of fact, the loss of the business as a whole.

The quality of products produced by a manufacturing concern depends to a large extent upon the quality attitude of all employes and their sincere desire to do the job right the first time. Recognizing that this is so, it is to the best interest of any concern to develop to the fullest a proper quality attitude on the part of all its employes. This is where the quality control people can do a great deal more to help stimulate quality interest.

At Harrison, our approach to this problem has been through improved communications with our employes and recognition of accomplishments made.

One lesson we have learned is that any program must be continuous. A one-shot program for 12 or 16 weeks will show improvement during the program; however, lasting effects can not be accomplished unless you continue with varied approaches from year to year. This is no different from any communication system. The advertising people have proven this beyond doubt with their constant, day in, day out, reminder of -

"the superlative 1960 Chevrolet"

"Pontiac - America's No. 1 road car"

"make yours an Olds in '60"

"Cadillac - for the man who cares"

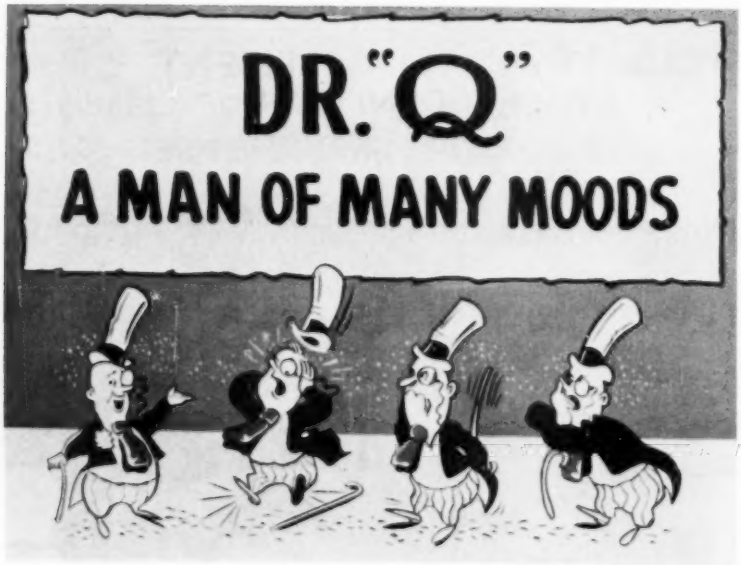
"the sheer look-Frigidaire"

"Harrison - temperatures made to order"

Through improved communications, we have told our quality story to all levels of management and to each of our hourly employes.

We have developed a poster series to illustrate the effect of quality on our customers and to dramatize various quality messages. These are effectively displayed at key locations in the plant.

Dr. "Q", and of course "Q" stands for Quality, has been our spokesman for the last four years. Dr. "Q" has been able to take an impartial viewpoint and express his opinions impersonally. This has been much more effective than communications from the Inspection Department.



Dr. "Q" has many moods---sometimes he is happy, sometimes he is mad, sometimes he is puzzled, and sometimes he is sad. These moods he transmits to our employes.

Lighted quality charts are placed in each department to show the departmental employes the quality level of the products they make. These charts are plotted daily and indicate whether we are improving or sliding backwards. They have been most effective.

Book matches are dispensed through the cigarette vending machines in the plant. They serve to convey a quality message to our employes. Each book contains twenty strikes for quality.

Additional publicity for our programs may be found in our local newspapers and in the Chronicle, our plant publication.

Product sample boards showing examples of good and poor quality items, as well as displays illustrating customer complaints, have been located in various departmental areas.

We have recently published two booklets, one entitled "Dial 'Q' for Quality" and the other "Reliability". In these booklets we have attempted to tell our employes our quality story--the importance of quality to us and to our customers, and the part each of us plays in producing quality products.

Our second approach to the problem has been through recognition for accomplishment. It is always easy to criticize for errors made; however, this is a negative approach and should be avoided. In recent years, we have attempted to take the positive approach and recognize good quality levels maintained and improvements accomplished. Dr. "Q" says "better your best."

Various quality awards have been made, such as:

Merchandise awards to employes making significant improvements in quality.

Pen and pencil sets for employes.

Quality plaques awarded to outstanding departments.

Quality plaques awarded for accomplishments by the Superintendents.

A quality banner awarded for plant accomplishment.

In fact, we have even awarded automobiles, based on divisional accomplishment.

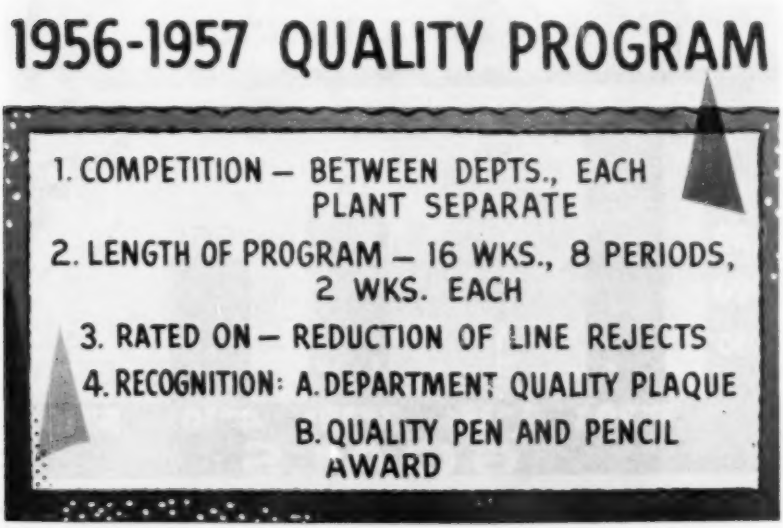
Monetary awards for suggestions leading to quality improvement have also been most effective.

Quality award luncheons have been held, at which time recognition has been given to various employes and departments.

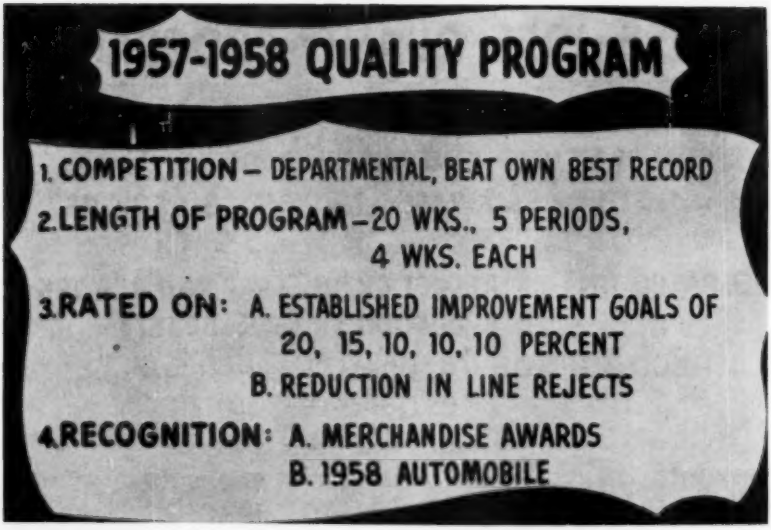
Our first program was held in the year 1955-56.

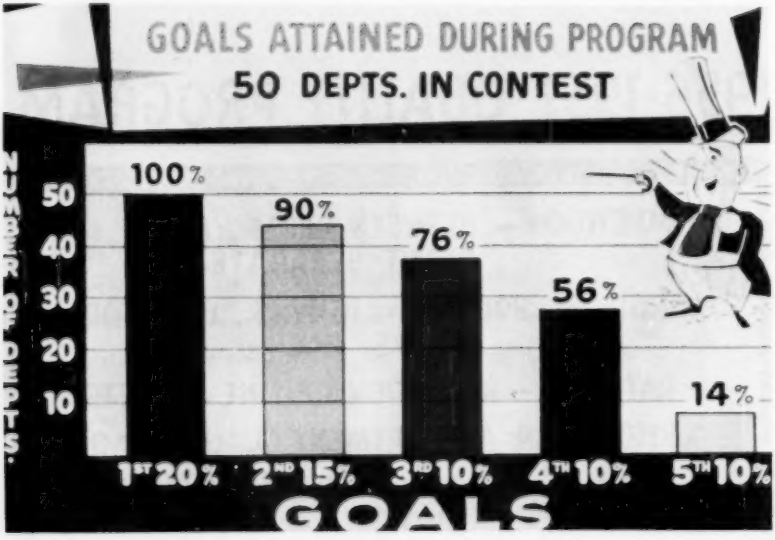
1955-1956 QUALITY PROGRAM
1. COMPETITION — BETWEEN PLANTS
2. LENGTH OF PROGRAM — 16 WKS., 8 PERIODS, 2 WKS. EACH
3. RATED ON — A. REDUCTION IN SCRAP AND REWORK
B. REDUCTION IN LINE REJECTS
4. RECOGNITION — A. QUALITY PENCIL
B. QUALITY BANNER
C. FOREMEN DINNER

In the year 1956-57 we altered our approach so that each plant had a separate program.


1956-1957 QUALITY PROGRAM
1. COMPETITION — BETWEEN DEPTS., EACH PLANT SEPARATE
2. LENGTH OF PROGRAM — 16 WKS., 8 PERIODS, 2 WKS. EACH
3. RATED ON — REDUCTION OF LINE REJECTS
4. RECOGNITION: A. DEPARTMENT QUALITY PLAQUE
B. QUALITY PEN AND PENCIL AWARD

Our theme for the 1957-58 program was "Quality Wins for You."


1957-1958 QUALITY PROGRAM
1. COMPETITION – DEPARTMENTAL, BEAT OWN BEST RECORD
2. LENGTH OF PROGRAM – 20 WKS., 5 PERIODS, 4 WKS. EACH
3. RATED ON: A. ESTABLISHED IMPROVEMENT GOALS OF 20, 15, 10, 10, 10 PERCENT
B. REDUCTION IN LINE REJECTS
4. RECOGNITION: A. MERCHANDISE AWARDS
B. 1958 AUTOMOBILE



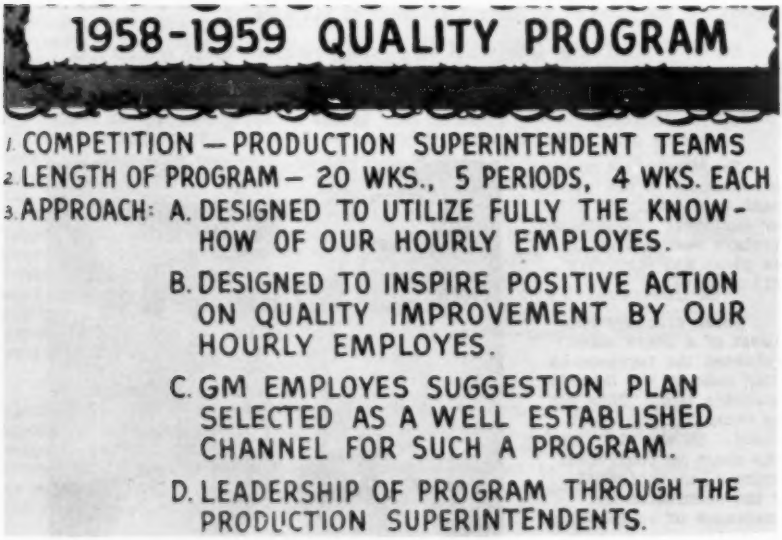
When you realize that these percentages of improvement were pyramided

after each period and the team was required to start from where they had finished the previous period, you can readily see the accomplishment which the shop made.

Our Scrap, Rework and Conditioning expense was reduced by 28.8%. The rejects coming to Inspection personnel were reduced by over 50% for the Division during this program. Our car heater plant reduced rejects over 70%.

In the year 1958-59, we felt there was a need for a new approach. We had used competition between plants. This was effective, but a little hard for the people to get too enthusiastic about--competing with employes in another plant whom they did not know. Competition between departments developed more employe interest, but had been used in our 1956-57 program. Our most recent program of beating your own best record in 1957-58 had been highly successful, but our evaluation of quality improvement had been confined to the reduction of scrap and rework costs and the reduction of rejects at our final inspection stations. We were looking for a program which would provide more stimulation to our supervisory group and also to our hourly employes.

Leadership and inspiration for the program came through the Production Superintendents' organization.


1958-1959 QUALITY PROGRAM
1. COMPETITION — PRODUCTION SUPERINTENDENT TEAMS
2. LENGTH OF PROGRAM — 20 WKS., 5 PERIODS, 4 WKS. EACH
3. APPROACH: A. DESIGNED TO UTILIZE FULLY THE KNOW-HOW OF OUR HOURLY EMPLOYES.
B. DESIGNED TO INSPIRE POSITIVE ACTION ON QUALITY IMPROVEMENT BY OUR HOURLY EMPLOYES.
C. GM EMPLOYES SUGGESTION PLAN SELECTED AS A WELL ESTABLISHED CHANNEL FOR SUCH A PROGRAM.
D. LEADERSHIP OF PROGRAM THROUGH THE PRODUCTION SUPERINTENDENTS.

To properly evaluate its effectiveness, we considered these points:

1. Improvement in scrap and rework costs.

2. Improvement in line rejects.

3. Participation in the Suggestion System by our hourly employes. Only suggestions for improving the quality of our products were considered. One point was awarded for each quality suggestion submitted, regardless of merit. Three points additional were awarded for each suggestion approved. This method of awarding points not only encouraged suggestions, but also awarded a premium for good quality suggestions. The Suggestion Rating was equal to:

$$\frac{\text{No. Submitted} + (\text{No. Accepted} \times 3)}{\text{No. Employes in Department}} \times 100$$

The Overall Quality Improvement Rating was equal to:

$$\frac{\text{Suggestion Rating} + \% \text{ Scrap \& Rework Improvement} + \% \text{ Line Reject Improvement}}{3}$$

The Superintendent's team showing the highest quality improvement rating for each period received recognition. A victory dinner and bowling party was held by the Production Superintendent with all members of his team attending. This included not only the Production Foremen, but also the people from Production Engineering, Process Engineering, Tool Engineering, Inspection, and Production Control assigned to his area.

Employes submitting acceptable suggestions received their recognition through the Suggestion Plan. Many suggestions were received during this period and a large percentage of them were acceptable to the Suggestion Award Committee.

Throughout the program, we continued to use various means to communicate with our employes. Illuminated displays were set up in the plant and were very well received.

These displays consisted of a chart which indicated the improvement being made by the Superintendent's team. This was the overall improvement record. To the right of this chart we publicized progress, as indicated by the submission and acceptance of suggestions for quality improvement. To the left of the chart we posted various quality messages, along with other information regarding the program.

Displays of good and poor quality workmanship were also most effective.

In our 1958-59 program, enthusiasm continued throughout the shop and this we believe was demonstrated by the fact that we had 5,512 suggestions submitted, of which 1,651 were Quality suggestions, or 29.9% of the total. We received an average of 83 Quality suggestions per week. During the period of this program, 650 Quality suggestions were adopted, or 39.5% of the total suggestions. In addition to the suggestion participation, line rejects coming to Inspection were reduced 33%. Scrap and Rework costs were reduced over 7%.

We are convinced that our programs have been well worthwhile in that each year the shop has made exceedingly good progress in reducing the number of rejects coming to Inspection. The importance of keeping good quality units coming to Inspection is self-evident. The customer is receiving more products built right the first time--and the Inspection group has less chance of missing rejects and letting them get shipped. This was accomplished even during the introduction of new products and processes into the plant. We changed from cellular to tube and center radiators; we added oil coolers to the radiators; and we took on a new business--automotive air conditioning.

It is true that we still find that, although we lose some ground each model change, our recovery time has been substantially reduced and the goals of last year are surpassed.

One of the most important benefits that we have gained from our programs is that the hourly employes and the team as a whole are more and more developing a strong quality attitude. As a matter of fact, employes are not hesitant to stop members of Management and discuss quality problems with them. They are quick to bring to Management's attention conditions which are preventing them from doing quality work.

There is one more point we should like to make. In our creation of the quality attitude in the minds of our employes, we at Harrison also gained something more--the employes' realization of the importance of the little things often overlooked in mass production operations. They learned that the little things could make the difference in winning or losing a contest, so they paid attention to small details. Because of our continuing quality improvement, we feel that this perhaps has become a habit with many of them. Because of their recognition of the importance of little things, they readily adapted this thinking to their participation in the quality improvement phase of the Suggestion Plan. They submitted little ideas as well as big ones, knowing that their suggestions, be they large or small, contributed importantly to the goal we continually had before us--complete customer satisfaction.

In conclusion, we must have determination not to "let well enough alone." We need to foresee the need for change and to act quickly and wisely to meet that need. Every person, whatever his job, must be motivated to perform reliably, for "good enough - isn't good enough." If every man and woman in every department of every plant is led and inspired to make Reliability his guiding star, then we will have made our Industry the most reliable organization of its kind in the world.

SOME ECONOMIC ASPECTS OF QUALITY STANDARDS

Eugene L. Grant and Lawrance F. Bell
Stanford University

"When I use a word," Humpty Dumpty said in a rather scornful tone, "it means just what I chose it to mean -- neither more nor less."

"The question is," said Alice, "whether you can make words mean so many different things."

"The question is," said Humpty Dumpty, "which is to be master -- that's all."

-- Lewis Carroll in "Through the Looking Glass"

Although the foregoing classic conversation has been quoted so often that it has lost all novelty, it continues to provide a concise statement of certain issues that arise in discussions regarding the meanings of words. The subject of our paper leads us into certain comments about the meanings attached to the words "quality" and "reliability." In our discussion we take a qualified position on the side of Humpty Dumpty; that is, it seems to us almost inevitable that when a popular word is used in a technical sense, it will have two or more meanings. The general thesis of our paper is that, in certain decisions on economic and administrative matters in industry, it is extremely important that the decision makers be conscious of the different meanings attached to the word "quality."

A year or so ago when we accepted the invitation to give this paper, our plan was to take a new look at certain aspects of a topic that was discussed by one of us at a national meeting of the American Society for Testing Materials here in San Francisco in 1949. In part, we have carried out this plan. The rather ambitious title of the 1949 paper was "The Economic Relationship between Design and Acceptance Specifications."[1]

In taking this new look at an old subject, we have found it essential to examine several different concepts associated with the word "quality." In part, therefore, the present paper has developed into an essay on the semantics of quality control.

This involvement in semantics has also involved us in a look at certain aspects of controversies that have arisen because of the current popularity of the word "reliability" and because of the distinctions that some persons have drawn between control of quality and control of reliability.

QUOTATIONS FROM A 1949 PAPER

The following three paragraphs quoted from the 1949 paper give a general idea of its subject matter. The first paragraph comes from near the start of the paper; the second comes from the middle; the third is the final paragraph summarizing the conclusions of the paper.

"Persons responsible for the preparation of specifications for manufactured products should recognize the distinction, first emphasized by Shewhart [2], between a design specification and an acceptance specification. The design specification deals with what is desired; the acceptance specification deals with the means of judging whether what is desired is actually obtained. My general thesis in this paper is that the setting of design and acceptance specifications should be viewed as two parts of a single problem of economy. This problem of economy cannot be solved without the point of view and the techniques of statistical method. It is in part a problem involving probabilities -- a problem in the economics of a calculated risk."

"One of the practical problems in establishing any sampling acceptance plan is created by the fact that no such plan can discriminate to any great extent among the lots from a statistically controlled process. This is true even of the best acceptance plans that can be devised. Even though many lots are rejected, the percentage of non-conforming product in the accepted lots from a controlled process will be about as great as the percentage in the submitted lots. The best that can be done in defense against a product statistically controlled at an unsatisfactory quality level is to reject such a high proportion of the submitted lots that it will no longer pay the manufacturer to submit the product. This might be accomplished entirely by the acceptance specification, or it might be accomplished in part by the introduction of a margin of safety into the design specification. The extent to which the design specification should be used for this purpose depends in part on the probable effect on the various costs."

"The main purpose of this paper has been to suggest that there will be real opportunities for cost savings and for improved quality assurance in industry if decisions on design specifications are no longer made in disregard of the acceptance specifications that are to be used. Generally speaking, this means that those responsible for decisions on the design specification must get together with those responsible for selecting and administering the acceptance specifications. If they are to get together to good purpose, both groups must acquire the point of view of statistical methods."

REACTIONS OF PRACTICAL QUALITY CONTROL MEN IN 1949

The foregoing quotations are not made with the purpose of calling anyone's attention to a paper that has doubtless been forgotten by almost everyone except its author. They were included primarily to provide a basis for commenting on certain reactions that were expressed to the author by some of his friends who had spent many years in administering quality assurance programs. A composite expression of these 1949 oral reactions might be phrased somewhat as follows:

"What you have said is correct enough in principle. We agree that it is true that designers often specify tolerances that are tighter than really needed for the satisfactory functioning of a manufactured product and that -- in effect -- such tolerances are often relaxed by means of the procedures used in acceptance inspection. Nevertheless, it is completely unrealistic to propose that persons employed in quality assurance ought to make suggestions to designers regarding design specifications; any such action would be resented by designers. Moreover, if we should be invited to make suggestions about design specifications, consistency would require us to invite designers to cooperate with us in establishing acceptance specifications. But we have learned from bitter experience that it is not a good plan to have designers making decisions about acceptance specifications. On the one hand, some designers are inclined to insist -- quite unrealistically -- that all product conform exactly to all design specifications; AQL's and AOQL's would all be zero per cent if some designers had their way. On the other hand, many designers are extremely unsophisticated on statistical matters and are inclined to place too much reliance on conclusions drawn from very small samples -- in many cases from a sample of one."

In their discussion in 1949, both the author of this ASTM paper and his friends working in quality assurance were assuming that designers normally had a clear idea of what they needed as aimed-at values of each quality characteristic; the main problems were those caused by the specification of tolerances.

The many problems associated with decisions regarding tolerances will always exist in mass production manufacturing; these problems existed long before 1949, they still arise in 1960, and they will continue to be troublesome in, say, 2049. But in certain other respects that affect the relationships between designers and persons concerned with quality assurance, the 1960 climate differs from the climate in 1949. One change from 1949 is that many designers are now quite sophisticated in their understanding of the concepts of statistics. Another change is that, particularly in the design of complex mechanisms, there is

widespread recognition that designers do not always know what ought to be the aimed-at values of quality characteristics of components; therefore, it is recognized that there is a need for information obtained from acceptance tests, from quality audits, and from end use of products to be fed back to designers. The rise of "reliability" as a popular word in the late 1950's was, in part, related to these two changes. Moreover, it seems to us that the historic barriers to communication between designers and persons in quality control have sometimes played a part in the setting up of so-called reliability departments that have, as one of their major functions, the duty of communication with designers.

SOME DIFFERENT CONCEPTS ASSOCIATED WITH THE WORD "QUALITY"

Shewhart's distinction between a design specification and an acceptance specification has already been mentioned. In a similar way, quality standards may be divided into standards dealing with quality of design and standards dealing with quality of conformance.

In applications to quality of conformance, the meaning of "quality" is simple and straightforward. Any failure of a manufactured product to conform to a design specification is technically a defect. Any product containing a defect is defective. From the viewpoint of quality of conformance, defective product is unsatisfactory; product free from defects is satisfactory.

In relationship to quality of design, a quality characteristic is any attribute of a component or finished product that may be specified or described. However, the concept of satisfactory quality of design is not as clear and definite as the concept of satisfactory quality of conformance. A person judging whether or not the quality of design of a complex product is satisfactory to him needs to have in mind what he believes to be the desirable attributes of that product, all things considered (including cost). Different persons therefore may have widely differing views of satisfactory quality of design of any particular product.

In the technical problems of judging satisfactory quality of design of components of a complex product, the judgment is related to the desired attributes of the complex product. Such judgments may depend even more on the appropriateness of the aimed-at values of the various characteristics of the components than on the tolerance limits that may be specified for these characteristics.

When the word "quality" is used in everyday speech with reference to manufactured product, it usually relates to the speaker's opinion regarding the general excellence of the product. This popular usage often combines the concepts of quality of design and quality of conformance. For example, a statement that the employees of a manufacturing plant have "qualitymindedness" may imply that they are proud of the quality of design of their product and also that they make every effort to achieve quality of conformance.

The following two examples are introduced to dramatize the difference between quality of design and quality of conformance. Both are true stories, only slightly altered to disguise the source.

EXAMPLE ONE

A housewife purchased a high-priced electric range made by the X Company, a well-known large manufacturer. This range had two ovens. Normally, the housewife used only the smaller oven. However, after some months of ownership of the range, she needed to use the large oven to cook a turkey. The pan containing the turkey rested on a grating supported by projections on the two walls of the oven. The housewife was extremely unhappy when the grating, turkey, and pan fell to the floor of the oven because the inner wall of the oven was not stiff enough to support its side of the loaded grating.

When she complained to the retailer who sold her the oven, he explained that the X Company had discovered this weakness in this particular model of range and had provided a kit to be installed without charge. The retailer's service man installed this kit. However, several months later when she again attempted to cook a turkey, she discovered that the kit had not been effective;

the turkey, pan, and grating again fell to the bottom of the oven.

This time the retailer explained that the guarantee period had expired and an alteration of the range could not be made without charge. However, he said that the company had now discovered the change that was needed to increase the stiffness of the inner wall of the large oven. When the service man came to make his estimate of the cost of this alteration (about $50), the housewife spoke quite unfavorably of the X Company. "Well," said the service man, "the X Company is usually pretty good but it sure pulled a boo-boo when it designed the oven for this range!"

EXAMPLE TWO

In a fairly complex electrical product for household use, there were many company designed parts that the Y Company purchased from outside vendors. For one such part purchased from Vendor A, the Y Company's designer specified an aimed-at value of a certain mechanical quality characteristic with tolerances of plus or minus ten per cent. Perhaps because special apparatus was needed to test this quality characteristic, no tests of this characteristic were made by the Y Company's receiving department. The parts purchased from Vendor A were installed in the complex product and functioned satisfactorily.

After this state of affairs had continued for several years, a few of these parts were tested as one of the operations in a quality audit. When none of these parts fell within specification limits, the parts in stock and the incoming lots were tested. The group making the audit finally reached the opinion that it was likely that not one of the many components purchased from Vendor A over a period of years had ever fallen within the specification limits. Moreover, the quality audit indicated that the finished product would have functioned less satisfactorily if Vendor A actually had made this part in accordance with the designer's specifications.

SOME POINTS IN SEMANTICS AND ECONOMICS ILLUSTRATED BY THE TWO EXAMPLES

Even if all the ranges of the particular model discussed in Example One had shown perfect quality of conformance to specifications, the ranges would not have satisfied their purchasers. The "boo-boo" made by the designer caused their quality of design to be unsatisfactory as judged by any reasonable standard. (Incidentally, the housewife did not pay the $50 demanded by the retailer to correct the designer's error. In effect, she redesigned the support for the grating in the large oven by using four empty cans to hold it up at its four corners.)

In contrast to Example One, the components of the complex product made by Vendor A in Example were 100% defective in the sense of showing a complete lack of conformance to specifications. Moreover, the quality of design was somewhat poorer than might reasonably have been expected. Nevertheless, in the sense in which "quality" ordinarily is used in everyday speech, the quality of the finished product was satisfactory. Surprisingly, the failure to check on quality of conformance offset the inadequacies in quality of design.

The examples emphasize the point that anyone who makes statements about the quality of a product needs to be clear about the meaning that he attaches to the word "quality." Moreover, it is not sufficient for a Humpty-Dumpty to be master over a word that he uses in several different senses; if he is to be understood, his listeners must be aware of the different meanings that he employs and must know when he shifts from one meaning to another.

The probable adverse economic consequences to Company X of the design error in Example One are evident even though it would have been difficult or impossible to measure them. There was surely a substantial loss of customer good will and confidence in the company's products because of the unfavorable reactions of the buyers of this particular model of electric range.

The economic aspects of Example Two are not quite so obvious. Company Y's avoidance of the costs of acceptance inspection did not have an adverse effect

on the product quality as judged by its customers. However, certain aspects of the case (not mentioned in the brief statement of the example) indicate that Company Y's costs may have been increased by its inappropriate design specification coupled with its absence of inspection. Vendor A, operating in a highly competitive industry, was consistently the low bidder for this component. Competing vendors based their bids on the assumption that the specifications would be enforced. If the bids of competing vendors had been based on components similar to those made by Vendor A rather than on the design specifications, their bids would doubtless have been considerably lower and the price paid for the components by Company Y might have been reduced.

SOME COMMENTS ON "RELIABILITY" AND THE SEMANTICS OF QUALITY

This paper has been written by two professors. In general, professors may be expected to observe developments in industry as spectators rather than as active participants. The following comments about reliability are made from this spectators' viewpoint; we do not feel that we are emotionally involved in any of the current controversies.

It is clear that "reliability" has become a magic word in American industry during the past few years -- particularly so in the defense industries. Moreover, in its modern sense, it is a relatively new word. The new magic of this word is not fully accounted for by the official definition of reliability as "the probability of no failure throughout a prescribed operating period."(3) There is nothing new about the need to specify the desired performance of certain types of manufactured product over a period of time under stated conditions of service; such performance is only one of many types of quality characteristics that have been specified by designers for many years.

As we have observed the duties of persons in reliability departments or reliability groups, it has seemed to us that an important part of their activity has been communication to designers. They have been concerned with the feedback to designers of information gained from tests and from observation of end-use of complex products. Their concern with quality of design has been even more with aimed-at values of quality characteristics than with tolerance limits for such characteristics. Part of their function has been to assist in avoiding designers' "boo-boos" on new product as well as to suggest design changes on existing product.

At the beginning of this paper, we mentioned the 1949 reactions of persons in quality assurance to the suggestion that design specifications and acceptance specifications should be viewed as interrelated matters to be discussed jointly by persons in design and persons in quality assurance. Their adverse reaction to this suggestion was understandable at the time. However, we cannot avoid the feeling that the development of reliability groups separate from quality control groups has been related -- at least in part -- to the barriers to communication that traditionally have existed between persons in design and persons in quality assurance.

Whatever organization structure a manufacturer may elect to use in trying to achieve desired quality of his product -- using the word "quality" in its nontechnical or popular sense -- it seems to us important that there be widespread recognition throughout a manufacturing organization that quality of design and quality of conformance are different matters and also that these different matters are closely related to one another. It seems to us that, in principle, the quality assurance function should deal with both quality of design and quality of conformance.

REFERENCES

1. Grant, E. L., "The Economic Relationship between Design and Acceptance Specifications," Symposium on Application of Statistics, Special Technical Publication No. 103, American Society for Testing Materials, Philadelphia, 1950.

2. Shewhart, W. A., "Some Aspects of Quality Control," Mechanical Engineering, vol. 56, December 1934, pp. 725-30.

3. "Reliability of Military Electronic Equipment," Report by Advisory Group on Reliability of Electronic Equipment, Office of the Assistant Secretary of Defense (Research and Engineering), U. S. Government Printing Office, Washington, D. C., 1957, p. 30.

QUALITY CONTROL PROGRAMS SHOULD BE COST REDUCTION PROGRAMS

Harmon S. Bayer, Quality Control Consultant
Bayer, Kobert and Associates, Detroit, Michigan

Few developments in the science of management more require sound definitions than the concept of the goal of a quality control program. Clearly, quality control programs cannot be considered effective unless they produce substantial cost savings for the organization. All too frequently this basic fact is not understood by those who administer these programs. Unfortunately, a large number of quality control managers still measure the effectiveness of their programs by the number of rejections of product, the number of reports and studies, and the use of other questionable criteria. It is not surprising, therefore, that many of these systems are considered by management to be wasteful. And as a result, the mortality rate of their managers is quite high.

It would perhaps be helpful, therefore, to consider step by step methods that management can employ to assure that a quality control program is, in fact, doing its job -- producing savings by reducing losses due to poor quality.

THE "COST" OF POOR QUALITY CONTROL

In organizing a cost reduction program for the quality control department the first requirement must be to assess losses that the firm suffers due to quality factors.

Here are some standard ones:

1. Scrap or waste produced in the production of the product.
2. Repair, rework or reoperation.
3. Downgraded products (seconds).
4. Inspection labor costs.
5. Customer service and rebates (payments to customers to settle quality claims, warranty cost, and policy cost).
6. Customer complaints (sometimes hard to dollar value, but loss of quality reputation translates itself into loss of sales).
7. Overtime hours caused by reruns or repairs of rejected material.
8. Down time caused by interrupted schedules required to remake rejected parts.

Often, obtaining these cost estimates is a difficult task since many organizations do not keep adequate records to enable proper assessment of their losses. In these cases, it may be necessary to make preliminary approximations of these losses. We know, of course, that a reasonable approximation can be obtained by the practical method of merely observing various departments for a period of time and recording the losses incurred during the observation period.

Surprisingly, many companies that have realistically appraised these and other quality costs have found that they can run as high as 10 to 50 percent of the entire manufacturing cost.

Thus, it becomes quite evident that these demonstrable losses can only be mitigated by a quality control program which, in fact, considers cost reduction as its primary target.

SELECTING QUALITY CONTROL PERSONNEL

A program which will produce substantial cost reductions logically requires mature, professional, competent personnel. Unfortunately, many companies attempt to institute quality control programs with unqualified personnel.

A quality control program should be administered by a manager who fully appreciates the need for: (a) competent inspection management, (b) a quality

control engineering staff with adequate experience in statistical engineering and statistical quality control.

This manager's staff, whether company personnel assigned to a new program or persons drawn from outside, should be motivated to develop skills in statistical quality control. Standards now being developed by the Professionalism Committee of the American Society for Quality Control may be used as a yardstick for performance that can be expected of these persons.

ORGANIZING THE QUALITY CONTROL -- COST REDUCTION JOB

To assure an effective program, it is necessary that the director of quality control adopt the attitude, "Anything that adversely affects quality is my business." Surely, this is not meant to imply that the quality control manager should unrealistically assume everyone's responsibility. But if any action by any department causes quality problems, it is up to the quality control department to aggressively seek out the facts and encourage those responsible to remove causes of poor quality.

With this in mind, the step by step approach toward quality control -- cost reduction programs is outlined in Figure 1:

(1) Establish goal. A program must have direction. The losses from poor quality listed above should be assessed and a goal and schedule established for reducing this total. Timidity at this point is not commensurate with the experiences in many companies that have had successful programs. Often a reduction goal of 50 percent has proved to be conservative. Many companies have shown 60 to 75 percent reductions over a five-year period. Subsidiary goals for achieving organization of systems, development of engineering staffs, and other needed planning should also be set at this time. But the fundamental target should be cost reductions.

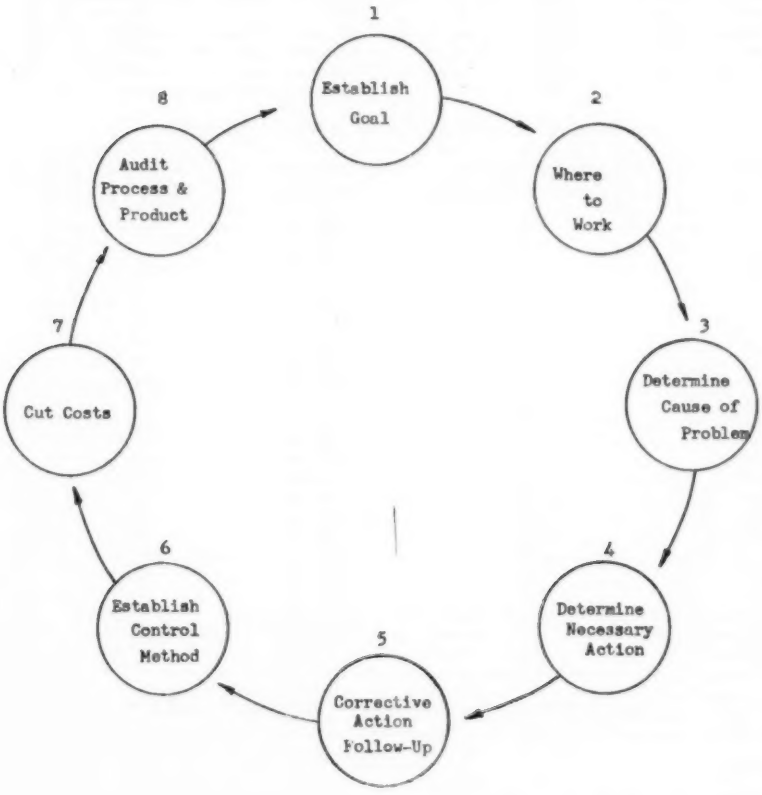
(2) Where to work. Usually most people expend their effort in unproductive activity, resulting in a minimum of tangible results. It is because they fail to understand that they have picked the wrong priority for their effort. A concept called Pareto's Principle of Maldistribution can be used to organize information to logically decide what areas of the problem are most susceptible to improvement. As a typical illustration of the principle, note Table II. Here are listed 23 production operators and the scrap each individual produced during one day. This is the usual way in which data is reported by an accounting system -- in this case, in order, according to clock number. To determine the overall pattern of these individual losses, these data are regrouped and placed in rank order from the smallest to the largest loss, retaining the identity of the source of the losses (see Table III). The third column in Table III is now cumlated from the bottom of the column (for example, 38.66+ 27.68 = 66.34; 66.34 + 20.58 = 86.92; etc.). The fourth column converts these data to percent cumulated.

Now we are ready to understand the real significance of our data. The column of figures in Table II is factual, but does not highlight, as does Table III and Figure 4, the knowledge that 5 out of 23 operators in the department were responsible for 79 percent of all the scrap. In this instance, the figures are presented for daily scrap only for ease in presentation. Actually it was found that monthly figures consistently showed individual operators in the approximate order that the daily figures showed.

If scrap and other losses are analyzed by these methods, one will determine that the following principle holds almost universally. A small number of areas where losses might occur will usually account for the major portion of those losses. For example:

3 departments	out of	10	will cause	approximately	80-90%	of	all	losses
3 machines	"	"	10	"	"	"	80-90%	"	"	"
3 people	"	"	10	"	"	"	80-90%	"	"	"
3 defect types	"	"	10	"	"	"	80-90%	"	"	"

FIGURE 1 -- ORGANIZING THE QUALITY CONTROL -- COST REDUCTION JOB



3 product identifications out of 10 will cause approx. 80-90% of all losses
Any 3 subdivisions " " 10 " " " 80-90% " " "

Taking into consideration this principle, one realizes that if a control program tries to blanket all areas, all departments, all machines, all people, it results in largely wasted effort. A much more efficient way to attack any problem is to develop a reporting system which identifies losses and reveals their responsibility in terms of departments, machines, operators, part identification, defect types, and any other logical subdivision. Identifying the sources of the losses and then concentrating on those few responsible areas will consequently reduce the effort drastically.

Admittedly, establishing the source information is extremely difficult. An accounting data processing system with a source document which indicates the cost center, identified to such factors as the department, person causing problem, machine, product identification and defect type is of inestimable value. Good cost reduction systems start here and modern data processing systems can quickly subgroup information into any desired breakdown. But

TABLE II -- SCRAP LOSSES BY CLOCK NUMBER

0003 -- 1.48	0038 -- 1.71	0068 -- 1.86	0118 -- 5.84
0004 -- 1.00	0039 -- 2.02	0072 -- 1.28	0130 -- 17.15
0005 -- .09	0047 -- 1.10	0077 -- .52	0146 -- 20.58
0006 -- 4.46	0054 -- 38.66	0083 -- .90	1176 -- .06
0011 -- 27.68	0062 -- .10	0085 -- 4.82	1465 -- .15
0031 -- .99	0066 -- 6.13	0117 -- 1.07	

TABLE III -- CALCULATING THE PARETO DISTRIBUTION

Clock Number	Scrap Losses Revised According to Rank Order	Rank Order Cumulated	Percent Rank Order Cumulated
1176	.06	139.65	100.00
0005	.09	139.59	99.96
0062	.10	139.50	99.89
1465	.15	139.40	99.82
0077	.52	139.25	99.71
0083	.90	138.73	99.34
0031	.99	137.83	98.70
0004	1.00	136.84	97.99
0117	1.07	135.84	97.27
0047	1.10	134.77	96.51
0072	1.28	133.67	95.72
0003	1.48	132.39	94.80
0038	1.71	130.91	93.74
0068	1.86	129.20	92.51
0039	2.02	127.34	91.18
0006	4.46	125.32	89.73
0085	4.82	120.86	86.54
0118	5.84	116.04	83.09
0066	6.13	110.20	78.91
0130	17.15	104.07	74.54
0146	20.58	86.92	62.24
0011	27.68	66.34	47.50
0054	38.66	38.66	27.68

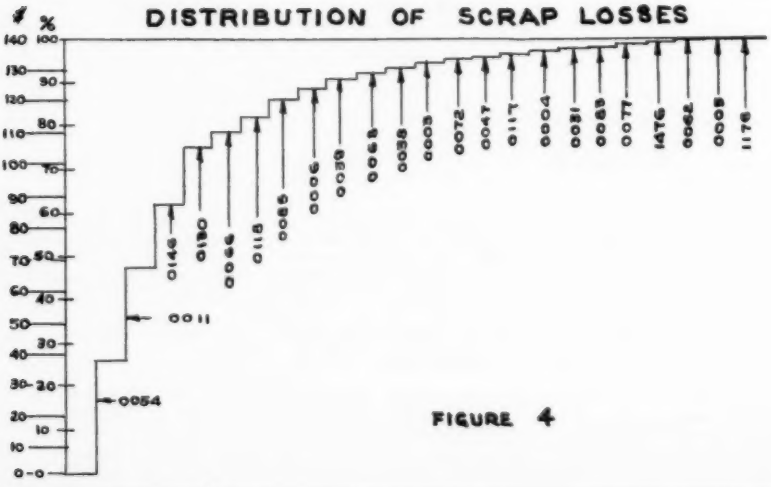


FIGURE 4

companies without modern data processing equipment should not think that they cannot develop similar systems. The author has proved that clerical time spent in analysis of such data is well worth the time and produces actual savings.

The manager responsible for the quality control department program should make certain that this kind of analysis is done on quality losses to assure that a minimum of effort is wasted on unnecessary areas, and effective stress is applied to the high loss areas.

(3) Determine cause of problem. The director of a quality control program should be aware that a true quality control program is characterized by defect prevention. The first step in defect prevention in any program is to establish a workable quality engineering effort. The working tools of this staff should be machine and process capability analysis, regression and correlation, design of experimentation, and other advanced statistical engineering methods. Above all, the staff should be well grounded in economic analysis methods so that their recommendations can be implemented by the dollar impact on the organization, and logical decisions can be made by management on their recommendations.

Through studies and reports generated by this group, the cause of the problem should be isolated and recommendations made for corrective action. Many organizations have staff which has traditionally carried out this responsibility. In instituting a new quality control program it is well to remember this. Quality analysis and statistical engineering services do not necessarily have to be within the quality control department. It is frequently more sensible, in the light of accepted practice, to reorient the existing analysis departments to quality engineering and statistical engineering techniques of analysis. The crux of the problem is that it is not generic whether the quality control department actually does the job, but that it is done by someone. In the event that it is done by another department, the quality control responsibility is to assure that it does get done by that department which has this responsibility.

(4) Determine necessary action. Too often programs fail, because after studies have been made, no concrete action is taken. The necessary action may entail a large expenditure and may cause upheaval of processes and equipment. The point to be stressed here is that assignments for action must be made. Assignment schedules must be projected to completion. If the action necessary is not planned and executed properly, the results can only be dissipated effort. The quality control director must accept a personal obligation to see that action and schedules are properly assigned by those responsible.

(5) Corrective action follow-up. Of all the techniques discussed in this paper, corrective action follow-up is by far the most important. It is a luxury to take for granted the notion that everyone will do his job. The successful administrator learns that constant follow-up is necessary to assure completed action.

As an example, let us take the following case and merely list the follow-up actions that were necessary to assure the removal of the largest scrap loss in one company. This was in a machining operation department. The quality control engineer in whose area the scrap was produced found it necessary to take these follow-up measures to assure successful completion of his project -- drastic reduction of the scrap:

(a) The original machine capability study determined that a size dimension was responsible for the scrap. A follow-up study of the scrap report after the size dimension was brought under control revealed only a slight lowering of scrap level. A follow-up study revealed that taper was the real cause, not the size dimension involved in the first study.

(b) A new tooling and fixturing method was established to reduce the scrap. Follow-up by the quality control engineer revealed that the tooling engineer, contrary to his assignment, failed to order new tools and fixtures

for the entire area. His reason -- his budget had been exceeded. The potential savings was better than 20 times the cost of tools and fixtures. A meeting with the tool engineer supervisor was necessary to get authority to purchase one set of experimental tools and fixtures.

(c) The purchasing agent, in turn, delayed the purchase of the tooling because of absenteeism of his stenographer. The quality control engineer convinced the purchasing agent this delay was deleterious to the entire program and it then swiftly was removed.

(d) The experimental study carried out by the production engineering department would have determined that the experimental method provided no improvement were it not for the fact that the quality control engineer observed the study and pointed out faulty control of the experiment. A restudy proved that the new tooling method was sound.

(e) Due to the expense of purchasing sufficient tools and fixtures for the entire area, the project now required a special appropriation request to top management. A three week delay was prevented because the quality engineer was able to convince the secretary of the top management committee to convene a special meeting rather than wait for the usual monthly meeting.

(f) Constant follow-up with the purchasing agent by the quality control engineer prevented a three month delay in the delivery of the special tools and fixtures.

(g) When the new tools were first used, a small improvement in scrap loss was noted. Follow-up by the quality control engineer led to a field engineer's visit by the tool manufacturer. Working with the quality control engineer and the production engineer, the field engineer from the tool manufacturer designed a simple but special adapter for the tools for the majority of machines in the department which were a different make from the one used in the original test. Thus it was shown that the tools were actually adequate.

(h) However, no substantial improvement was noted in the scrap loss reduction. The quality control engineer found that the operators were unfamiliar with the new tools and special training was needed. After the training of the operators, the scrap figure finally took a drastic drop.

(i) However, this improved scrap figure did not remain low. A further study by the quality control engineer noted that some operators were going back to the old type tool. It was discovered that the new tool gave the operators a slightly lower production rate, adversely affecting incentive payments. A restudy of the operations by the industrial engineering department determined a basic methods change involving the new tools which increased the production rate to compare favorably with the previous production rate, and therefore did not adversely affect employee's incentive payments.

(j) Constant follow-up has been necessary to assure that old type tooling will not be used by those operators who still prefer to use the old type tooling.

One must realize that the above case is typical for many operations. The reason for high losses is certainly varied. Reducing these losses is not always simple. What must be understood is that failure to take any one of the follow-up steps described above could have neutralized previous efforts if the quality control engineer did not have the tenacity to see his job through.

In many instances, engineers with a variety of technical backgrounds place undue emphasis on solving management problems by merely using formulae and studies. They frequently dismiss the human and organizational problems as unimportant or easily accomplished. The effective manager on today's industrial scene is aware that it is only through such follow-up as has been described above, by determined individuals, that results are obtained.

(6) Establish control method. Determining the basic technical problem is one thing. Getting corrective action is another. Both of them depend upon the establishing of a control method that can signal trouble. Some types of control methods necessary are as follows:

(a) Quality loss reporting and analysis systems daily.
(b) In-process control chart systems.
(c) Operator and inspector checking systems to assure good product.
(d) In-process and final acceptance inspection systems.

Of course, each of these methods will establish the need for current studies, action and follow-up.

(7) Cutting the cost of the control system. Many control systems start with effort placed in logical areas. But as the control system spreads, the effort engulfs the available control manpower. Managers fail to understand that, once solved, the causes of a particular problem may have disappeared. However, we find continually that routine habits become ingrained and people continue to place unnecessary effort in these routines. They fail to realize that during the solving of the problem, people have been trained, new tools and equipment have been purchased, and new production skills have developed. It may now not be necessary to concentrate as much effort on the problem itself. Effective system managers must learn to reduce routine checking and effort by stretching out frequency of checks or dropping them entirely to release manpower for the next problem in order of importance. Of course, if the reduction of effort is too drastic for the problem and the problem recurs, the system must be flexible enough to allow re-establishing the necessary effort to control the situation.

(8) Audit process and product. To assure that the problem continues to be in control, an effective audit system should be established. Customer complaints must be checked regularly. In-process material and finish product must be checked occasionally for the specific defect which has been brought under control. As experience demonstrates the ability of the organization to keep the characteristic in control, the frequency of audit may be reduced drastically. Often the control audit may not be on the product but on the process -- for example: A strength problem in a chemical process is now adequately controlled by merely checking a dryer temperature which has proved to be an effective barometer of the strength. The dryer shows excellent control when the system is operating correctly, but drifts when an unbalance develops in the system. A daily check of the dryer has proved sufficient to prevent undesirable characteristics in the process, which, when in control, produces product with adequate strength.

Of course, if the audit indicates that a problem is recurring, it may be necessary to retrace some of the various steps in the "loop" as shown on Figure 1. However, good experience files should prevent retracing studies previously completed.

By following this "loop" we now have an effective way of removing problems in logical sequence.

WHO IS RESPONSIBLE FOR QUALITY

It is the author's opinion that the best way to establish an effective quality control program is to start with the concept that the operator is fully responsible for the quality of his product, and the ultimate goal is to give him every technique to carry out his responsibility. The same concept sees inspection as merely a "crutch" which the operator will continuously lean upon. Therefore, the author believes the full responsibility for quality must be given to the operator, and his foreman, and systems must be developed which throw the entire responsibility to the operator and gradually removes the inspection "crutch."

Here is how the system may be installed:

(1) The quality capabilities of all operations and processes causing quality difficulties are studied to assure that management is not responsible for the causing of poor quality. It is not logical to expect an operator to produce quality which is over and above that which the process or machine can produce. This is a management responsibility. The causes of poor quality must then be identified and removed to assure that the operator is now fully capable of maintaining the quality of his operation. Only when this is done can management fairly require that the operator be fully responsible for his quality.

(2) The operator is given all the information he needs about what is expected from him in terms of quality, specifications, etc. This is done by developing a production quality instruction sheet for each operation. This is usually a switch from the past -- all too frequently we find operators working in the dark as to the quality requirements necessary. Often the inspector is better informed than the operator. Sometimes both are confused.

(3) The operator is given all the gages and technical equipment he needs to check enough product to assure himself that the quality is adequate. We frequently find inspection supervisors objecting to this; the excuse is that the operator is not to be trusted with delicate gages. This is usually "empire building" on the part of inspection personnel. Consider the fact that we trust an operator with production equipment whose cost far exceeds that of any gage. It is seldom that one finds an operator who cannot be trained to use any gage. Sometimes it is impossible to give an operator a gage because of the length of time needed to check a characteristic.

(4) The operator is given sufficient time in his pay standard to check enough pieces to assure himself that he is making adequate quality. Upon investigation, we may find that an excess of checking time is already in his standard and it costs no more to gain adequate control. However, it may be necessary to change the standard to include some inspection time. Then we can fairly require quality of the operator, since we now know he has the opportunity to observe for himself the quality of his product. In the event the operator must have a special checking method, such as a control chart or a sampling plan, the operator is paid to perform the charting or sampling for himself to help make him aware of the problem of adequate quality.

(5) An adequate in-process identification system is developed so that any product can be attributed to the operator performing each step of the production. When each operator is finished with the product, he signs an "operator certification ticket" (usually the lot identification card) and this signifies that he certifies to the quality of the product he produced.

(6) Inspectors are assigned to check product at each stage of the operation -- only checking certified lots to determine whether each individual operator can, in fact, certify his lots. Reports are made to production supervisors on all operators and their attention is called to chronic "unacceptable certifiable operators."

(7) "Certifiable operator" is made a condition of employment and corrective action is taken with any chronic cases of operators who are not performing adequately, of course, authority for this action rests with the production supervisor.

(8) When it becomes apparent that rejects and scrap and quality are improved, the inspectors are gradually pulled from the process and from the final line. Sufficient audit inspections are retained to assure that the quality picture does not later develop an undesirable turn. Usually final inspection can be curtailed or even removed entirely except on those performance characteristics which are of a critical or legal nature and must be checked.

SCHEDULE YOUR PROGRAM

One principle of good management which is too often ignored by the quality

control department is to establish a specific schedule for the program. Goals must be set for activities and for the actual dollar pay-off of the program. Subgoals should be assigned to supervisors in the program so that the entire program coordinates toward meeting projected over all goals.

One word of caution. Don't be unrealistic in your plans. A full, effective program usually takes about 5 years. However, the actual breakpoint of savings overcoming expenditures should materialize between 1-2 years after the program is launched. From that point on, savings should increase.

One other note. Schedules are made to be met -- not adjusted to agree with what has actually been accomplished. The difference between the effective manager and the ineffective one is the ability of the former to plan, schedule and produce.

TYPICAL EXAMPLES OF RESULTS

(1) Within 9 months a large brass manufacturer saved in excess of $100,000 over and above the entire investment in his quality control program. His customer complaints were drastically reduced.

(2) A rubber manufacturer paid off the investment in his program and produced further savings by reducing his scrap losses to less than 20 percent of the former level within 6 months. Customer rebates were also drastically reduced.

(3) Through the quality control program a high precision machining plant revealed some basic problems of machines and methods. These were corrected and quality improved. Notwithstanding the cost of correcting the processes, the program was paying for itself within four months after it was started. Customer reaction improved steadily while overall savings mounted.

(4) A ceramic manufacturer has substantially reduced his scrap level. His quality complaints have been reduced to a minimum. The quality control engineering department has become most effective in working with the organization to achieve: (a) the constant improvement of process methods, (b) substantial production cost curtailment.

CONCLUSION

Briefly, then, how does one establish a quality control -- cost reduction program?

First, the concept is established within the organization that the operator and his foreman are fully responsible for the quality of his operation.

Second, systems are developed to assure that those responsible for poor quality (management or operator) are identified.

Third, the economic order of each loss is evaluated and priorities for action are chosen (Figure 4).

Fourth, each priority problem is attacked with the methods illustrated in Figure 1. One by one the "loop" shown is closed on each problem and the result is a sharp reduction in losses in scrap, repair, inspection labor, customer complaints and rebates and other quality losses.

MECHANIZATION, QUALITY CONTROL, AND AUTOMATION

by
Kenneth W. Davidson, Norman S. Ince, A. W. Wortham

INTRODUCTION

In the rapid growing field of electronics, the problems associated with increasing quality and new demands for reliability represent an ever increasing challenge. There is a direct relationship between the emphasis placed on the reliability and quality of a product, and the importance of the functions which they perform.

Increasing quantities of life tests and environmental tests to measure and prove reliability as well as grade quality, place an increasing demand for automation, mechanization, and revised Quality Control procedures.

This paper presents the need of the research into new concepts of automatic test facilities which will provide the types of equipment that are required to constantly evaluate, in quantity the reliability of electronic systems and components.

The need of a facility designed to provide highly reliable and precise test data for transistors and other miniaturized electronic components by automatically testing and recording parameter values is discussed.

THEORETICAL RELATIONSHIP OF THE UNCONTROLLED VARIABLES TO THE HOMOGENEITY OF DEVICES

The American consumer has raised his quality standards to a new high and the military customer has been confronted with very stringent reliability requirements. Industry in general is as concerned with quality assurance as the American public or the Government. Responsible manufacturers are keenly aware of the need for reliable industrial and military equipment, and in some instances have set up extensive programs to assure its attainment.

Mechanization and automation have in the past years become keys to the success of many types of industry. The rapid pace set by the technology progress in the electronics industry has voided standard quality control methods. Many of the quality control practices which are routine for tried and proven industries must be cast aside for the more theoretical and untried methods whereby many of the new problems confronting the electronics industry may be solved.

The ideal situation in the component manufacturers plant would be good incoming materials, excellent process fabrication techniques with properly balanced quality control, adequate technical and semi skilled personnel, high volume production with automation, low cost high quality end-product and satisfied customers. However, many of these attributes are complicated with detering factors such as:

1) Meager concepts of quality control due to lack of competent personnel
2) New devices introduced without proper test equipment to measure and evaluate them
3) Changes in techniques of production and modified production processes creating a need for means of rapidly evaluating the new methods prior to mass production
4) Large volume production created by product demand without sufficient time to analyze information obtained on production problems

5) Keeping abreast of the industry requirements causing new devices to be developed and produced without sufficient technical information on end applications
6) Unproven processes during early "state of the art" developments creating the need for a method of quickly choosing, adapting and controlling the right process for meeting the customer's needs.

METHODS OF CONTROLLING THE VARIABLES

To alleviate the detering factors previously discussed, several interim steps may be persued. Complete automation is usually an ideal situation although it is not always desirable. Management must consider all aspects of the steps which will lead them to the solution of their specific problems. Figure 1 shows a closed loop system of material flow and communication links in a typical component manufacturing process. Beginning at the incoming materials stage there are several types of measurements which are performed on a wide variety of raw materials. These types of measurements include chemical, electrical, mechanical and thermal. In a majority of the cases this link does not readily lend itself to complete automation or mechanization.

The in-process fabrication link is easier to adapt to automation and mechanization because of the following factors:

1) Whereas in the incoming materials stage the wide variety of materials are checked and analyzed for use in many different processes, the materials used on specific production lines are more homogenous, thus making a more uniform product
2) The repetitive nature of specific points within the manufacturing process which may be mechanized and measured.

In the semiconductor industry the final acceptance link is most

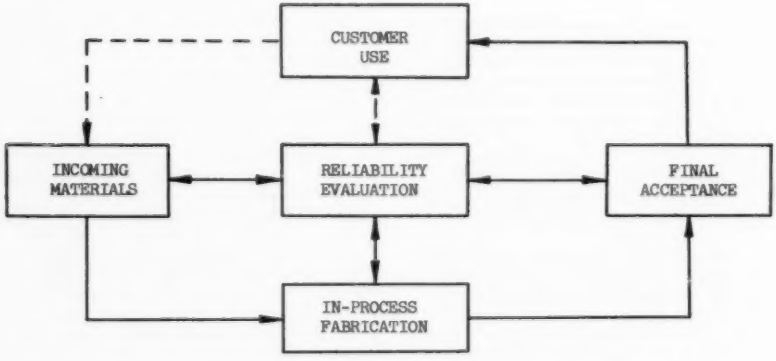


Figure 1
Material Flow and Communication Links in a Typical Component Manufacturing Process

susceptible to complete automation and mechanization for many reasons. Among those are:

1) Many tests and measurements must be performed on large quantities of like devices
2) Electrical classification and mechanical sorting must be accomplished on every device made.

In every instance where large volumes of devices are subjected to 100% testing, the use of automatic equipment is a prime requirement to ensure acceptable results.

During the early stages of a new industry such as semiconductor manufacturing, large quantity measurements of the various parameters of the products were done by hand. Special hand test sets had to be set up for different types of device evaluations, and distributions of the data were made by desk clerks. Since most of the readings taken were visual and the data recorded was by hand, many chances for error were introduced.

As production requirements increased the brute force hand testing and measuring methods had to be replaced. Automatic testing methods offered increased speed, greater accuracy, improved safety, reduced costs, greater flexibility, and as an option, a permanent data record.

One of the innovations in the final acceptance link by Texas Instruments Semiconductor Components Division was the Centralized Automatic Tester (CAT). This facility is a semiautomatic transistor testing machine suitable for production use. Figure 2 shows the loading end of the machine and the main control console. The control console supplies all of the power and timing signals to operate and cycle the test and sorting stations. The test rate can be adjusted so that up to 2000 units per hour can be processed.

Figure 2

Figure 3

A temporary memory in the form of a punched card is carried along with each transistor as it is moved down the test line. Sorting is done from the information punched on the card with a mask and a photocell. Acceptable units and selected rejects are automatically removed from the machines and placed in bins. Figure 3 shows some of the test and sorting stations. The machine is semiautomatic in that the units must be loaded by hand into test blocks. To assure the accuracy and repeatability of the results of CAT tests, a special equipment control system has been designed and put into operation by the Quality Assurance Department.

Test Standards

As the CAT machine is a "go, no-go" type operation and, as such, no variables data is available; the type of control is of a "go, no-go" nature. Transistors are used as the test standards. These devices have been chosen to represent the complete parameter distribution of a given family, including rejects. This sample is tested manually on accurate equipment so that it may be determined in advance which transistor will be sorted into the various bins. Devices are then color coded according to the appropriate bin.

Test Sequence

The flow chart of Figure 4 shows the various steps taken during the control procedure. The test blocks are permanently marked for positive identification and the number of both the transistor and the test block are written on the punch card which serves as the memory unit. As the block moves along the line, various punches are made in the card at the test stations; the number of holes depends on whether or not the transistor meets the particular limit at each of the test stations. When a given unit has been tested it is pulled into one of the sorting bins at the end of the line. The exact bin in which the test unit

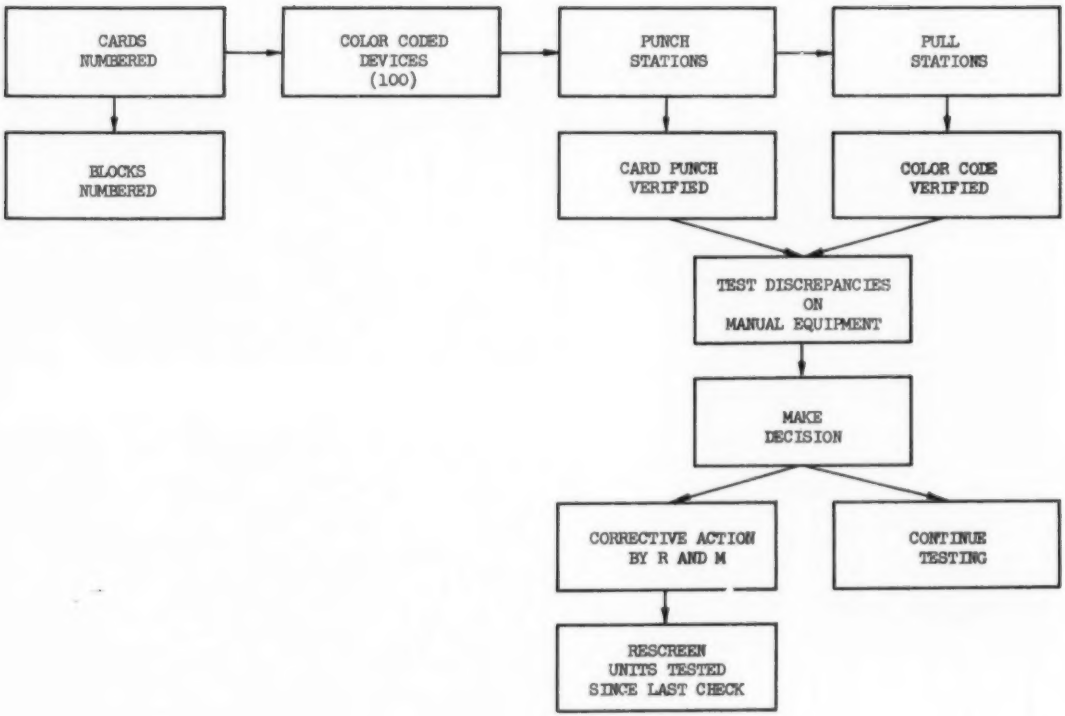

CARDS NUMBERED
COLOR CODED DEVICES (100)
PUNCH STATIONS
PULL STATIONS
BLOCKS NUMBERED
CARD PUNCH VERIFIED
COLOR CODE VERIFIED
TEST DISCREPANCIES ON MANUAL EQUIPMENT
MAKE DECISION
CORRECTIVE ACTION BY R AND M
CONTINUE TESTING
RESCREEN UNITS TESTED SINCE LAST CHECK

Figure 4

appears depends upon the arrangement of holes punched in the card. Since each transistor has been checked in advance for each parameter the CAT is sorting, it should be known that the card will assume a definite hole pattern and the transistor will end up in a certain bin.

Analysis and Corrective Action

At the end of the test line the punch cards are collected and held for further reference. When all of the units have been sorted into bins, the technician makes a visual inspection of all the bins looking for a color-coded unit which might be in the wrong bin. When one is located, the card for that device is checked to locate the error. When this is determined, the unit is tested on a hand test set to ascertain whether or not the CAT machine was in error. When the CAT machine is found to be in error, necessary corrective action is taken by qualified personnel. A further check on the accuracy is then made by comparing the punch cards to a set of cards that have been previously punched according to each unit's classification. If any difference in the punches is noted, it is then determined which punch station is in error. After the necessary corrective action is taken the units which were in error are retested as a final check. Occasionally, punches are made on a card and, upon checking the stations in error, no apparent reason for the error can be determined. Then the block number is taken from the card and the block is checked to determine if it is defective. When a block is found to be defective it is removed from the test line until it has been repaired.

Other Tests

There are other control systems that could be used that would be quite satisfactory, such as sampling each bin on a regular basis to spot errors. However, this technique requires that each transistor of the sample be tested for every test station that appears on the CAT. By the time results of a test of this nature are completed the CAT may have tested several thousand transistors in the interim.

Advantages

The main advantages of the system presently employed are:

1) Since the same transistors are used each day, an analysis of the repeatability of the CAT is readily available
2) Fast feedback of information is possible
3) Minimum personnel is required to conduct the tests
4) Complete analysis of the punch and pull stations permits maintenance to locate trouble spots quickly.

Keeping in stride with Engineering technology in the automation field the Facility for Automatic Sorting and Testing (FAST) machine was developed. One of the basic advantages of this machine is that feedback information in the form of variables data is available.

Figure 5 shows an overall view of the machine. The individual test stations in this equipment contain automatic recalibrating circuits and additional provisions to turn the equipment off should a portion malfunction. Each test station is provided with a readout system so that data may be recorded on punched tape or typewritten. Another feature of the equipment is that the transistor is placed in one socket and is passed through all tests before being removed. This method keeps the lead flexing during testing to an absolute minimum.

A third and final example whereby automation is essential is that of life testing for Reliability Evaluation. To provide the services needed

Figure 5

for Reliability Evaluation on a large scale operation, complete automation is highly desirable. The facilities to provide such a service must incorporate:

1) Variables data in useable form
2) High reliability of test
3) High resolution
4) High sensitivity
5) Maximum accuracy consistent with the state of the art
6) Methods whereby special tests above and below normal temperatures may be run.

The Sequential Mechanism for Automatic Recording and Testing (SMART) is such an automatic test facility which was designed and built at Texas Instruments for use in Reliability Evaluation. Figure 6 shows the SMART machine. It is a machine capable of performing up to 16 sequential tests on a device while recording the desired data in a three-digit readout on a punched card. Accessory consoles can be connected to extend this range to cover AC, high frequency, and pulse tests.

The sequencing is at the control of the operator in a Sequence Control Unit. Any number of tests between 1 and 16 may be selected or skipped. If only a few tests are planned, these tests may be performed and the Sequence Control is homed without stepping through the remaining tests. The conditions of any test may be selected on a Program Unit. The variables of programming include level and point of application of test voltages or currents; point of grounding test circuit; duration of test; level; polarity; terminals where readout is to be measured; and level of reject trips. The readout is fed to a differential input digital voltmeter. The digital voltmeter is used with an external trigger, supplied by a timer circuit. A digital display of the readout is provided on the front panel. A transfer register also converts the output into relay closures. These closures are in turn used to activate punches in an IBM card punch.

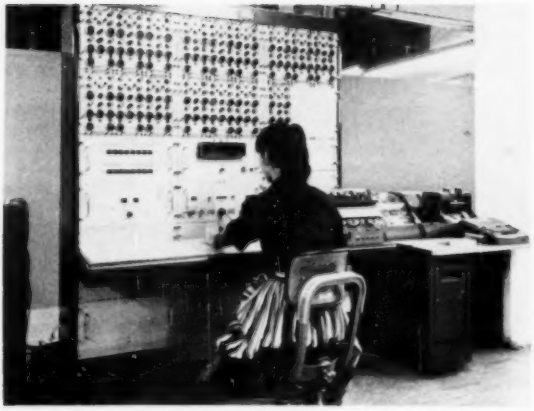


Figure 6

Once the data has been accumulated it becomes necessary to perform standard statistical analyses and special engineering evaluations. The statistical analyses may be readily done using computer programs.

CONCLUSIONS

The increasing quality demands and higher reliability requirements in the Semiconductor Industry can only be met by automation. Some outstanding benefits derived from automation include:

1) Greater Accuracy
   a. Correlation factors are established to provide methods of prediction so that the standards continually represent the product quality which the customer desires
   b. Selection of the parts at the proper quality level can be made
2) Increased Flexibility
   A rapid and continuous check into the effects of materials used on the deterioration of the finished parts with time is made
3) Permanent Data Record
   It provides management with sufficient information to make decisions in regard to cost versus market.

The manufacturer of Semiconductor Components must provide increased mechanization and automation:

1) To adequately provide feedback within his own plant
2) To provide increased customer services as required
3) To meet the increased high reliability requirements placed on his products.

The manufacturer must also refine the basic quality control concepts to keep abreast of the changes which occur due to automation.

COMPANY STANDARDS AND QUALITY CONTROL

Major General Leslie E. Simon (Ret.), Staff Director of Research
and
Charles A. Bicking, Manager, Quality Control Branch
Research and Development Division

The Carborundum Company
Niagara Falls, New York

Introduction

Research and development work in corporations is tending to weld with all other technology into a continuous spectrum of applied science.

Statistical Quality Control has a brilliant opportunity to contribute to this constructive change by becoming the principal moderator of applied science throughout the corporation.

The principal instruments for moderating technology are company standards. The preparation of good standards is a formidable task because it requires sharp attention to statistical detail. The essentials of statistical methodology requisite to preparation or application of standards should exist in the corporation at three levels: (1) the staff of the corporation, (2) the staff of each manufacturing division, and (3) the production line.

Quality Control as the Moderator of Applied Science

Many of today's most successful corporations use applied statistics in a variety of ways ranging from process control and quality assurance to design of experiment and operations research. Whereas names for various phases of these statistical activities differ, much is to be said for embracing all industrially applied statistics under the well accepted term quality control, both because of its acceptance and because the preponderance of the activities are aimed at quality control. Looked upon in this fashion, quality control is recognizable as a part of the scientific method, which has been defined as repeated cycles in the continuum of (1) hypothesis, (2) experiment, and (3) judgment of hypothesis.

The whole of corporate management follows an analogous process consisting of (1) generation of plans and policies, (2) conduct of business or manufacture in conformance with the plans, and (3) inspection or appraisal of results to determine the degree of success in reaching goals. Phrasing this brief description of the scientific method in the terminology of statistical quality control, the place of quality control is (1) in the formulation of plans that are valid; (2) in manufacture that produces satisfactory quality with optimum profits; and (3) in the devising of inspection and appraisal techniques that are just, reliable, and effective.

The emergence of the leverage power of modern statistical quality control in applied science is very similar to the birth and subsequent development of the powerful modern scientific methodology known as operations research. As a matter of fact,

one might even say that the former was a precursor of, and an indispensable adjunct to, the latter.

At about the turn of the century, the evolution of new things (research and development, if you please) lost its simplicity of the zero order with the virtual demise of the envied and vilified lone inventor. The concept of group enterprise in research and development was a great stimulus to fruitfulness, even though the group inventors were all of the same discipline. Whereas the need for persons of different disciplines was sensed, from time to time, the inhibiting effect of the lack of consciously planned diversity of disciplines was not realized until World War II was upon us. World War II brought into sharp focus the efficacy of diversifying the team through a "mingling of the disciplines." In practice, operations research brought out the fact that the member or members of the applied science team who represented the discipline of mathematics generally emerged as the directing agency of the joint efforts, irrespective of whether a boss, project leader, or chief investigator was formally designated. Perhaps this is not surprising, if we remember the words of the famous Harvard philosopher, C. I. Lewis, "Of all human endeavors, mathematics admits most of certitude and precision."

In practical work that involves diverse disciplines, that somewhat humbler branch of mathematics, applied statistics, is represented in almost every well-planned group enterprise, whether or not the person who exercises the statistical functions professes to be a statistician. Thus, the statistician, either through merit or faute de mieux, automatically appears to become the agency for reducing what would be chaos to orderly simplicity. He often detects the necessity for the enterprise; he prompts the planning; he designs the experiment; he separates the spurious from the real, the noise from the signal, the strong parameters from the weak; his logic leads to proper task assignments to components of the team; and he analyzes the data and presents the results. In short, he is the moderator.

Reliance of Company Standards on Sound Specifications

Among the principal instruments for moderating technology are company standards. Company standards are based on sound specifications. The preparation of specifications, their organization and their application are, therefore, among the most important activities of quality control.

A standard may be defined thus: "A standard is a specification accepted by recognized authority as the most practical and appropriate current solution of a recurring problem." (1)

In a similar vein, a specification is defined as follows: "A specification is a concise statement of the requirements for a material, process, method, procedure, or service, including the exact procedure by which it can be determined that the conditions are met within the tolerances specified in the statement."(1) From the viewpoint of quality control, the important thing about the specification is not what it describes but that it consists of "a concise statement of the requirements" and that it includes "the exact procedures by which it can be determined that the conditions are met within the tolerances specified."

Throughout the following discussion, the term "specification" is used, with the understanding that a specification recognized by authority and in recurrent use is, in fact, a standard.

We can employ any classification of specifications that we wish, and organize them as we see fit. The object in mind in establishing a sound system of specifications should be to make sure of three things:

1- That the requirements of the specifications are clearly and completely set forth;

2- That the specifications contain, or reference, the procedure by which it may be verified that the requirements have been met;

3- That, if a specification be standardized for general use, it be approved by management and accepted by the consensus of its users.

Insofar as is possible, every document in a complete system of specifications should possess these three important elements.

For purposes of control, it is probably a good idea to classify specifications in a small number of distinct groups. One basic grouping is:

1- Product specifications
2- Process specifications
3- Drawings
4- Inspection specifications

Subclassifications may be established as convenient. For example:

1- Product specifications
   a- Raw material specifications (or purchase specifications)
   b- In-process product specifications
   c- Final (or end) product specifications

2- Process specifications
   a- Manufacturing specifications (or manufacturing methods, descriptions of manufacture, etc.)
   b- Marking and packing specifications

3- Drawings, as applicable

4- Inspection specifications (or procedures)
   a- Test specifications (or methods)
   b- Sampling specifications (or plans)
   c- Control specifications (or procedures)
   d- Quality audit specifications (or procedures)
   e- Standardization and calibration specifications (or gauge control procedure)
   f- Preventive maintenance specifications (or procedures)

Specifications may be of a temporary nature or may incorporate special modifications for particular customers. They may eventually become standards and be so designated. It is desirable

to have some way of marking each such document to indicate its status.

The preparation of good specifications is a formidable task because it requires:

1- Sharp attention to statistical detail;

2- Definition of quality not only in physical terms, but also in terms of acceptable level and limits of permissable variation therefrom;

3- Provision for verification of attainment of the specified quality.

It is exactly in the above respects that many specifications are deficient. They often state only desired levels with no expressed limits. When limits are specified, they are often arbitrary and are not developed from nor consistent with process capabilities. The lack of limits or the imposition of unreasonable limits leads directly to inconsistent application of the specifications. Individual judgment or unwritten rules are applied to make exceptions and develop a grey area, or area of indifference, in which grade or other product classifications become meaningless.

Unless definite limits are applied and rules for conforming to the limits are adhered to, the quality achieved can not be known with any certainty. Merely aiming for a goal does not provide a guarantee that it has been met nor that the scatter about the aiming point will be satisfactory to the user of the product.

Often, even if limits are applied, it is not essential to know precisely the confidence with which they have been met. For example, we may determine 95% process capability limits and be satisfied if all subsequent lots of the product fall within the limits. Or, a specific level of assurance may be written into the inspection procedure which may spell out carefully how large the sample shall be, how it shall be obtained, and how it shall be treated and tested. On the other hand, it is becoming more and more frequent to specify the confidence required in so many numbers. Thus, it may be specified that it be demonstrated with a given confidence, say 95%, that not less than a given percentage, say 99%, of the lot will be within stated limits. It is important that the specification either imply or state the extent to which non-conformance with the specification will be tolerated.

Another common lack in specifications, is clear direction of what to do if the specification is not met. The simplest instruction is, of course, to reject the lot in question. In some circumstances, however, it may be desirable to resample or to increase the sample size. Possible courses of action on the lot may include further processing, reprocessing, blending, or accepting it as a lower grade.

Setting Levels and Limits in Specifications

Statistical methods may be used to establish the specification limits in the first instance. There are numerous references in quality control literature to the statistical bases of specifications. (See References 2 through 6). Whether the object is setting up a new specification or reviewing an old one, the statistical procedure is the same. The first step is to determine what the production process is capable of doing. It is then possible to determine whether the process performs within that

capability consistently. This is done best by means of a control chart study.

The control chart is not the only means of arriving at the natural limits of a process. In the past, engineering judgment has been relied upon successfully in many instances. In the absence of adequate data, judgment may still have to be relied upon. Judgment relies upon some kind of data, or experience. The better the experience is quantified, the sounder the basis of judgment. Eventually the degree of quantification exceeds the ability of the engineer to establish judgments without the aid of statistical analysis. This is where the quality control engineer with his understanding of distribution theory, control charts, and experiment design steps into the picture.

If great quantities of data are already available, the quality control engineer knows how to analyze them to establish limits of confidence. If the data have to be gathered from a going production process, there is no special need to depart from accepted control chart techniques for computing limits. If the process is still in the experimental or the pilot plant stage, it may be desirable to gather the data in accordance with a designed experiment.

Whatever the data source, natural or inherent limits of the process are used to establish achievable, economic limits in the specification.

Setting Verification Procedures in Specifications

In some instances, absolute limits in specifications may be desirable. Each lot must be compared to a rigidly fixed limit and be rejected if it does not comply. More often, however, there is implied or expressed, some acceptable level of compliance. We are satisfied if the evidence shows that there is a specified degree of confidence that the limits are complied with. This confidence may be expressed with respect to either individual items within a lot or with respect to its average quality.

As examples of established procedures which may be drawn upon to verify that the quality specified is achieved with a known level of confidence, there are sampling plans which apply to any industrial condition. As has been mentioned earlier, the requirement may be for a certification statement. In this event, the method of collecting the data and of analyzing results may not be indicated. Any available appropriate statistical techniques of experiment design and analysis may be used to secure the data on which the certification statement is based.

The Organizational Place of Quality Control

The need for the existence of capabilities to handle statistical methodology at various levels in a corporation makes it desirable to discuss the quality control organization in some detail.

A need for an organization for quality control exists at three levels in the corporation: (1) in the central staff level where the most advanced types of work are done, such as operations research, complex design of experiment, and corporationwide assessment of level of product quality; (2) in the technical branch level where there is the statistical staff function for the decentralized manufacturing division that is merely a small-scale model of the statistical function that the central staff carries out for the whole corporation; and (3) in the manufacturing branch

level where the quality of manufactured product in the process of production is controlled in the Shewhart sense.

Thus, there is a clear channel of technical communication from persons concerned primarily with simple process control techniques such as control charts for $\bar{X}$ and R, all the way up to persons concerned with the most advanced statistical techniques including the programming of their work on electronic digital computers. No man need fear getting beyond his depth, for he can always get help from the next higher echelon. This is a vertical flow of technology between persons of like disciplines; and applies not only to quality control, but to all required disciplines.

Similarly, there is a horizontal flow of communication between quality control and the other disciplines of the same organizational echelon; e.g., between quality control branch of the central staff and the other staff branches; between the quality control department of a technical branch and its other departments; and between the process quality control engineer and the foreman, superintendent, or plant manager. It is primarily the consequences of the horizontal flow of communication that result in quality control's function as moderator of applied science, although the enunciation of the problem that requires the task group approach may be a complex interaction of vertical and horizontal communication.

One essential for good product quality control is that it be organized so as to exercise properly its functions and responsibilities with neither infringement on other fields of technology nor surrender of any of its powers to them. Necessarily, many organizational units are involved in the chain from top staff to the "pay-off" functions of manufacturing and sales. There must be a concise statement and a clear understanding of the functions and responsibilities of each organizational unit with respect to the role it plays in the several sequential phases of quality control. For example, the determination of what is proper quality and the establishment of tolerance limits on the acceptable variability therein must be distinguished clearly from the actual achievement of that quality level in production. The former is a setting of standards, and is exercised by the manager of the operating division through his technical branch. The latter cannot be separated from the general responsibility for manufacture; is specifically the domain of the manager of the manufacturing branch; and should include detailed control in the process of production. This distinction follows the sound principle that whoever is responsible for achieving proper quality in manufacture should not also be responsible for determining what constitutes the proper quality against which the achieved level will be judged.

Quality Control at the Central Staff Level

In the central staff of The Carborundum Company, the Research and Development Division is the counterpart of the technical branches of the operating divisions. As such, it has staff responsibility for product quality control throughout the Company, including:

(1) Technical assistance to the technical branches on specifications, processes and methods of control. This includes assistance in instrumentation and process automation, particularly with respect to new or improved methods of measurement and control.

(2) Performance of laboratory analyses and tests not

within the economic capabilities of the technical branches.

(3) Periodic audits of the performance of quality control within the operating divisions.

(4) Responsibility for approving product and process specifications prepared by producing divisions for products that are substantially used by another operating division. This responsibility extends to refereeing incompatibilities among the respective divisions concerning specifications of one which impinge on the quality characteristics of the products of another division, including the determination and issue of product and process specifications.

(5) Supplies all statistical services required by the Central Laboratories of the Company.

(6) Assisting in the training of quality control personnel in technical and manufacturing branches of the operating divisions, imparting to these personnel information on new developments in the field, and enhancing the professional capabilities of all those engaged in product quality control.

Quality Control at the Divisional Level

Within a manufacturing operating division there are three branches with substantial responsibilities in quality control. They are: The Technical Branch, The Manufacturing Branch, and The Sales Branch.

The Technical Branch of the Division

The technical branch has the responsibility of determining (a) what constitutes proper quality for materials purchased, including components and supplies; (b) what the quality of outgoing manufactured product should be; and (c) for determining whether the manufacturing branch is meeting standards by conducting independent quality audits. More specifically these responsibilities include:

(1) Determining what the significant quality characteristics of a product are and the extent to which it should have them. Such a responsibility requires seeing to it that the technical, manufacturing, and sales branches have a mutual understanding of the environments in which the products will be used, which, in effect, defines the nature of their commercial applications.

(2) Determining, preparing, and issuing product and process specifications, standard inspection procedures and amplifying documents, such as descriptions of manufacture. (This is a very sophisticated and exacting task).

(3) Determining, preparing, and issuing the methods and operational procedures to be used to check and control product quality including laboratory analysis, testing and inspection.

(4) Periodic auditing of the application of product and process specifications, and of testing and inspection methods within the manufacturing branch.

(5) Establishing and monitoring the use of a system for gauge control and for routine calibration of control, test and inspection instrumentation and equipment to keep it accurate and performing properly.

The Manufacturing Branch of the Division

The manufacturing branch has the responsibility of manufacturing products of the quality delineated by the product and process specifications and of exercising economic control of quality in the manufacturing process, in accordance with the methods and procedures prescribed by the technical branch. It can and should make recommendations for changes in such procedures to either lower costs or to improve quality, but it cannot unilaterally change or modify such procedures or substitute for them. The manufacturing branch:

(1) In responsible for elimination of losses.

(2) Has control over all equipment, inspectors, and others engaged in any way in quality control during the process of production in the manufacturing plant.

(3) Is responsible for the carrying out of statistical and other quality control procedures, including use of statistical techniques in sampling, maintaining process capabilities, controlling precision and accuracy of testing, and applying statistical process control, as prescribed by the technical branch.

(4) Is responsible for interrupting production if product quality is unsatisfactory, until remedial action is taken. (However, if the stoppage of production is because of faulty requirements, the remedial action is a responsibility of the technical branch, and it must also bear the onus for the interruption of production.)

(5) Is responsible for engineering action to identify and locate the assignable cause of trouble, when off-quality is indicated; and is responsible also for correcting the fault or for securing corrective action.

The Sales Branch of the Division

The sales branch is responsible for continuous supply of information from the field to the technical and manufacturing branches with respect to:

(1) The apparent relative quality of competitors' products.

(2) The needs of customers with respect to new and improved products and product quality and prospective new applications of Carborundum products.

(3) The apparent quality or reputation of the company products being shipped to the customers. It is also

responsible for obtaining the necessary support from the technical branch with respect to cooperative tests run in customers' plants.

Conclusion

In today's rapid pace of changing technology and the concomitant shortage of scientists and engineers, the West's holding its own depends as never before on utmost skill and efficiency in technology. We owe it to our way of life to see to it that the great leverage power of modern statistical quality control is brought into full play in the economic struggle.

References

1. Chingas, J. and Glie, R., "Definitions in Standardization", The Magazine of Standards, October 1959, pp. 304-6.

2. Shewhart, W.A., "Some Aspects of Quality Control", Mechanical Engineering, December 1934, pp. 725-730.

3. Dodge, H.F., "Discussion on Statistical Quality Control and Its Relation to Specifications", Discussion on Statistical Quality Control in Its Application to Specification Requirements, Pittsburgh Spring Meeting, American Society For Testing Materials, February 26, 1946, p. 1.

4. Goffman, Casper and Manuele, Joseph, "Use of Statistics in Writing Specifications", ASTM Bulletin, No. 139, March 1946.

5. Simon, L.E., "Dollars for Your Thoughts", ASTM Bulletin, No. 139, March 1946.

6. Bicking, C.A., "The Use of Statistical Methods in Determining Product Specifications", TAPPI, Vol. 32, No. 9, September 1949, pp. 405-9.

AUTOMATIC PRODUCTION RECORDING - YIELD - AND QUALITY CONTROL

James A. Curry
Kaiser Steel Corporation

In February, 1954, the Operating Management of Kaiser Steel Corporation inaugurated a formalized yield improvement program. A section within the Industrial Engineering Department was established and named the "Quality Control and Yield Group." This group was staffed with Industrial Engineering personnel who had had previous statistical and/or Statistical Quality Control experience.

At the same time a supervisor in the Iron and Steel Division of the Metallurgical Department was assigned to represent the Metallurgical Department in this yield improvement program.

Partially by circumstance, but mostly because it seemed a logical place to start, most of the yield improvement effort for the first year was concentrated in the primary rolling mill.

The Industrial Engineer, the Metallurgist, and the Superintendent organized as a yield committee for the primary mill. Within a very short time certain types of avoidable losses became apparent. Within a year the overall yield in this mill was increased about one per cent. These gains were accomplished by a variety of means; some involved methods changes, some involved changes in steel ordering practice, and still others involved closer supervision of operator performance which was aided by charts and reports supplied to operating supervision on a current basis.

It soon became apparent, however, that variation, although possibly not the root of all evil, was at least responsible for most, if not all, of the avoidable losses which were being experienced.

To cite an example, an ingot has a nominal weight of 18,700 pounds, from which the primary mill is supposed to get three slabs each 100 inches long, six inches thick and 32 inches wide. These slab sizes have been calculated by Mill Scheduling to give the correct amount of steel from which the Plate Mill is to make a plate of a given dimension and weight within close tolerances. If this slab is too light, then the Plate Mill will be unable to produce the right size plate. On the other hand, if the slab is too heavy, the Plate Mill will have excessive scrap losses from shearing the sides and ends.

To carry the example forward, the steel making unit pours ingots for a plate order. The Plate Mill rolls the slabs produced and 5% of them do not make the ordered dimension and/or weight. Why? There are, of course, innumerable possible reasons, among which the following would be some examples:

1. The primary mill roller might have produced the slabs less than 32" wide or less than 6" thick.

2. The primary mill shearman might have sheared some of the slabs too short.

3. The Plate Mill roller might have rolled too thick or too wide; in either case, this would probably result in a short plate.

4. Possibly there was not 18,700 pounds of steel in the ingot to start with because of the variation in pouring in the Open Hearth.

5. Perhaps there was an abnormal amount of "pipe" in the ingot so that the shearman had to cut back the top of the ingot an abnormal amount in order to cut out the unsound steel.

These are only a few of the possible sources of variation which could account for the undesirable end result--the ordered plate was not produced.

On the other hand, if the ingots or the slabs were too heavy, this results in an excess yield loss too, which is also a highly undesirable result, since good sound steel would be returned to the Open Hearth as scrap which should have gone into finished product.

To isolate these individual causes of variation in an integrated steel plant without interfering with production is an impossible situation unless there are integrated into the production process some check points which, at that time, we did not have.

Starting with a theoretical ingot weight, we went to a theoretical slab weight and then to an actual plate weight and dimension.

How much do ingots vary from nominal weight? It is true that we do have an ingot scale on which we weighed each ingot buggy. Each buggy carried from two to four ingots depending on size of ingot. From this, it might have been possible to determine individual ingot variation from the variation experienced in groups of ingots, but even here we found these results unreliable because, we found a bias in the scale weights, which made any estimate of individual variation very unreliable. For instance, if in weighing over this scale the drag was stopped as each successive buggy was centered on the scale, one weight was obtained, but if the scaleman tried to get a "flying weight," then an entirely different weight was obtained, depending on how he operated the scale. Then too we found that the amount and direction of this bias was dependent on how fast the drag was moving over the scales. In fact, subsequent investigation revealed that the theoretical ingot weights were actually more reliable than the scale weights were under the conditions which we were weighing.

It was possible to take ingots from various sized molds from the ingot yard and weigh them on a truck scale. From this we were able to make an approximation of the variability of ingots which were nominally the same size. But, considering that each unit weighs from five to fifteen tons, developing enough data to make a reliable estimate represents a real material handling problem, both from the standpoint of the mechanics of getting it done as well as from the economic considerations of cost.

We did, however, develop a program wherein enough ingots of each size were weighed to develop the pattern of variation with a good degree of confidence. We found, generally speaking, that individual ingots varied in weight in the range of ± 5%. We were also able to isolate the two major variation sources, the foundry where the molds were made, and the Open Hearth pouring practice. We concluded that about 3% of this variation developed because of variation in mold volume and about 2% originated in the Open Hearth. These two sources were of course additive which accounted for the 5% total.

As for the slabs which were rolled and cut from these ingots, here was an entirely different source of variation. Rolling slabs from ingots to a predetermined width and thickness has not as yet, at our plant, become an exact science. What were the capabilities as far as both the equipment and the operators were concerned? This too had to be established. Here again we were working with theoretical weights which were developed from ordered dimensions.

The approach of measuring slab length, thickness and width was tried; but, with the exception of length, we came to the decision that measurements as we were able to take them on a production basis, were not a very good index of the amount of steel in the slab. Here again it was apparent that before any real constructive work leading to a reduction in variation in the amount of steel in the slabs could be accomplished, it would be necessary to be able to determine what the variation was and whether it developed primarily from cross section variation or length variation. Weighing of slabs by transporting them to a truck scale, although not as much of a problem from a material handling viewpoint as weighing ingots, was still practically impossible without interfering with routine operating practices.

These then, in brief, are some of the conditions encountered which led us to the conclusion that, in order to realize the potential yield improvement which we knew existed in the primary mill area, it would be necessary to develop some means of obtaining both ingot and slab weights on a day-to-day basis and to use these weights as a method of isolating and fixing responsibility for the excessive variability which we were experiencing.

In 1954 and 1955 contact was made with suppliers of electronic scale equipment and a proposal was developed for the installation of a scale into the primary mill approach table which would weigh ingots just prior to reduction into slabs or blooms. A second scale would be integrated into the discharge conveyor table which could weigh the slabs and blooms produced. It was reasoned that, by establishing the incoming weight and by weighing the product produced, not only could the extent of the variability be determined, but we could also develop standard yields which would be a good measure of performance of the mill operation.

Both Management and Operating Supervision were in full accord with the objectives but before approvals for the capital expenditures could be obtained, an expansion program involving the primary mills as well as other areas in the plant began to take form. It soon became apparent that a new primary mill facility was to be included as a part of the plant expansion and as a consequence this first scale proposal was shelved with the understanding that adequate weighing devices would be incorporated as an integral part of the new mill.

In the course of planning the expansion program, other groups became involved from the methods and procedure anbles of recording production. These groups reasoned that it was a logical step to proceed from the use of the ingot and slab scales which were originally conceived as a tool to be used on a sampling plan basis, to the concept of integrating them into a punched card system which would convert our present accounting system utilizing manual recording of theoretical charge and production figures to automatic recording of actual charge and production figures. This was, of course, far removed from the idea which the yield committee had originally proposed. But, since the whole cost and operation of the system could be justified on the basis of savings in accounting costs, the yield committee for the Primary Mill gave the program enthusiastic support.

As a result, specifications for the Automatic Production Recording system were incorporated as a part of the prime contract for the new mill.

The complete system as designed and furnished by two sub-contractors, one for the weighing devices and one for the recording and punched card equipment, consists of the following basic units:

1. A typewriter card punch which is located in the Mill Scheduling office. The primary mill rolling order is produced on this equipment as well as a punched card for each ingot on the order. All common information concerning heat, steel type, grade, ingot size, and theoretical weight is prepunched into each ingot card. The primary mill rolling order along with the prepunched cards are delivered to the mill recorder prior to rolling.

2. An electronic ingot weighing scale built into the ingot turn-around which weighs the ingot and transduces the values into impulses accepted by the recording equipment. This scale has a maximum capacity of 80,000 pounds divided into two range increments of 40,000 pounds each.

3. A program or control unit which includes a clock, calculating equipment to determine delay time, and keyboards and punch to insert variable type information into the prepunched ingot cards originally prepared in Mill Scheduling.

4. The slab recording system consists of an electronic scale of 40,000 pounds maximum capacity divided into two ranges; a program or control unit which includes calculating units to accumulate the totals of ingot and slab weight, switches and keyboards to insert variable information; a modified card reader used to read the ingot cards, a modified card-punch used to punch a card for each slab, duplicating the parent ingot data as well as slab weight, etc.; and finally an electric typewriter which prepares a written report of all slab activity and periodic totals of ingot and slab weights. This typewriter is tied into the slab recording system.

As designed, this system was developed to give complete accounting information for production, delays, incentive wage payments and, as a by-product, a gold mine of information for the yield committee which would enable us to isolate causes of variation in ingots as well as slabs. In addition, the fact that the information is available immediately, opens up all sorts of possibilities for control of variation. This, we believe, is the whole key to making real yield gains in this area.

I am well aware that it is almost a universal custom in papers of this type to pat yourself on the back, boast of your successes, and generally try to impress everyone concerned by enumerating the tremendous gains which have been made. This is a human trait which is almost universal. We at Kaiser Steel are no different in this respect than anyone else.

But, the fact is that we have very little in the way of successes to report on this project. When judged on the basis of its primary purpose--automatic production recording--the system has been a failure in that it has not produced an automatic production record which either the Accounting Department or the Yield Committee (except for closely supervised sampling runs) can accept. In fact, as presently designed, there is considerable doubt that it can ever produce sustained automatically recorded data which will serve the day by day requirements of either group or of Operations Management.

On the other hand, under certain limited conditions, it can and does produce information on individual ingot and slab weight variations which is valuable in a yield improvement program.

The reason the system has not been a success from an accounting standpoint is that the system has not yet worked for three successive turns without producing some bad figures. So it is not possible for them to take the totals produced by the system and apply them for accounting purposes. Since it is necessary to maintain a manual parallel production recording system, the Accounting Department would rather use their manually generated figures than to reconcile the errors and misses generated by the automatic recording system.

On the other hand, for sampling tests by qualified personnel, where the scale is known to be correctly calibrated both before and after the test is completed, accurate information can be developed on variation of ingot weights, variation of slab weights and yields; and, as a result of this information, we are well along in correcting some of the major sources of variation which is after all a major part of our objective. However, for this purpose alone, we now feel that less sophisticated--and more economical equipment--would suffice.

It may be profitable in a situation like this to enumerate the causes of failure and to analyze each one for the purpose of determining what needs to be done to correct it. I suppose this is essentially what the missile people do each time a rocket fails to perform in the expected manner. However, let me add parenthetically that I do not necessarily subscribe to their idea that, even though the missile failed, the experiment was a success.

I think the causes of failure can be divided into two main categories, i.e.;

1. Equipment failure.
2. Personnel failures.

Of these two, we feel that equipment failures are much more critical than personnel failures in this case.

It is believed by our instrument maintenance department at Kaiser Steel that the electronic load cell is not stable enough in its present state of development to be used as a routine weighing device. This is true not only of the electronic scales in the Automatic Production Recording system, but in other installations of electronic weighing equipment in our mill as well, some of which are located in much more favorable environments than the scales in the Slabbing Mill.

Such variables as ambient temperature, very small line voltage changes, shock loads, moisture, oil, vibration, all of these things, probably affect the read-out obtained from an electronic load cell. Under optimum environment conditions perhaps the load cell may perform well, but anyone familiar with steel mill conditions will appreciate that optimum environment as concerns the factors listed is impractical if not impossible to attain.

The only other major difficulty with the equipment was with the card reader. The function of this piece of equipment was to read out both the prepunched information and the variable information on the ingot card and to transmit this information to a card punch, which in turn punched a slab card for each slab produced from the parent ingot.

It perhaps should be explained at this point that the recording and punched card portions of the system were engineered by one division of the Contractor, which at that time operated almost as a separate organization. In engineering this system, they used some of their own components which they had designed and some standard components. The reader was from this division, and it has been admitted, at least unofficially by the parent company, that the reader used was not adequate and should not have been incorporated into the system. The malfunctioning of this reader had the effect of producing mis-punched cards. The magnitude of the errors introduced was so great that it was necessary to manually edit all cards prior to preparation of production and yield reports by Machine Accounting. Even after performing this manual edit, the undetected errors were so large as to make the resultant report unsuitable for accounting use.

In the meantime, there was an internal reorganization within the contracting company which transferred the maintenance of the system over to their Industrial Division. The reader was replaced by a modified standard reader. Since this change, the reader has practically ceased to be a problem.

In the introduction of this system, as with all departures from established routine, we anticipated that there would be many problems concerning the personnel involved. Operating personnel have to push switches to activate the system; Shearmen have to activate the weighing mechanism for slabs in addition to their regular duties; the Mill Recorder had to acquire an entirely new set of skills; the Order Writers in Mill Scheduling had to learn to use electric typewriters and operate a card punch where before, rolling orders were hand written. However, we were very pleased to find out that these problems turned out to be of less magnitude than we had expected. Our major trouble in this area has been in obtaining proper maintenance of the scales and mechanical transducing equipment.

It is undoubtedly true that the training of good electronic scale repair and maintenance men takes considerable time and since maintenance of this equipment is very different from instrument repair work regularly found in a steel mill, this fact has produced for us a real problem.

Another problem, and one which has not as yet been satisfactorily resolved, is the problem of zeroing in the scales. This problem develops of course in the inherent instability of the electronic load cell. In order to maintain calibration, it is necessary to put a test weight on the scale--25,000 pounds for the ingot scale and 4,000 pounds for the slab scale. So as not to interfere with mill production, this has to be done during a mill delay.

Since in most cases mill delays cannot be anticipated in advance, the instrument maintenance personnel are usually not readily available, and before they can be located and brought to the mill, the delay is over. On the other hand, to keep a maintenance man on standby to wait for a delay is an uneconomical practice also. However, if greater stability could be engineered into the weighing mechanism, this problem would largely disappear.

I would like to re-emphasize the fact, though, that the failures in both equipment and personnel are much more critical from the standpoint of accounting applications than they are from the point of view of their use to accumulate basic variability and performance data on a statistical sampling basis which is, after all, the original purpose which the quality control and yield committee had in mind. The reason for this is that accounting requirements for inventory, incentive pay, and cost analysis, require continuing accurate information, whereas statistical analysis requirements can be met with sample data obtained during intermittent, closely controlled periods.

What of the future as regards the application of this type of equipment to the steel industry? As far as we at Kaiser Steel are concerned, it is a somewhat muddied situation. We are now of the opinion that the system we have has serious shortcomings. The muddiness in the situation resolves itself around the question as to whether it has the capabilities of ever performing satisfactorily for accounting purposes. If not, then it is much too elaborate, and incidentally too costly a system to obtain information for yield and quality control purposes alone. It is somewhat like buying a fine precision jeweled watch to do a job for which a $3.00 alarm clock is better suited.

On the other hand, the Primary Mill Operating Management and Supervision have had a taste of what this type of installation can really do from the standpoint of yield and quality improvement, and are insistent that all ideas of discontinuance of the present system be considered only on the basis of replacement by some better and more economical means of obtaining essentially the same basic information.

The operating personnel have found this system to be very beneficial in maintaining the identity of the material through the rolling, shearing, and marking operations. Another benefit, and this may be the one most lucrative, is the use of the slab weights obtained as a process control device. During periods when the scale is known to be supplying accurate slab weights, it has been possible to detect changes or drifts in the rolling or shearing operations. Whenever the slabs are not cutting out as scheduled, an examination of the ingot and slab weights permits an immediate determination of whether this is due to insufficient ingot weight or due to rolling or shearing practices. Incorrect rolling or shearing practices can then be corrected immediately, thereby reducing yield losses.

What happens in this area in both the immediate future and the long range depends primarily upon resolution of the problem of insuring that a weighing device will consistently provide accurate weights of ingots and slabs under the conditions which prevail in the Slabbing Mill.

The ability to utilize reliable impulses from a weighing device to automatically enter the actual weights into data recording and processing equipment presents a relatively minor problem according to our systems personnel. But, as I have indicated previously, without accurate weights initially, no amount of subsequent sophisticated mechanized or electronic data processing will make the information usable where day-by-day reliability is necessary.

There is little disagreement with the desirability of automating and mechanizing the recording and processing of production data when it can be done accurately and economically. The problem is to keep objectives and systems in step with the "state of the art." It appears that our present system may have been developed ahead of the "state of the art."

Our maintenance and repair people feel that from a practical day-by-day operating viewpoint, the electronic load cell is an impractical device in the environment in which it is generally used in a steel mill. If this is true, and our experience would seem to lend credence to this idea, then regardless of any computer and recording limitations, the system is unacceptable for production recording.

At present, investigation is being carried forward as follows:

1. The Engineering Department is carrying forward studies to determine the feasibility of substituting either mechanical scales or hydraulic scales for the existing electronic scale components.

2. In conjunction, the Data Systems Department is investigating the possibilities of a much less complex system which would give at least most of the basic information which we are now getting.

In conclusion then, it appears that the direction we will take in this program would possibly be to search out more reliable weighing methods, to devise a less sophisticated, but at least semi-automatic recording system--one which would adequately satisfy the requirements of both operating management and yield and quality control activities--and possibly, in the near future, depending on advances in data processing equipment reliability, again move toward the ultimate goal of automatic production recording.

SAMPLING FROZEN BEEF USING A VARIABLE SAMPLING PLAN FOR PERCENT FAT

Gerald J. Desmond, Quality Control Manager
Gerber Products Company, Oakland, California

ABSTRACT

A variable sampling plan for acceptance sampling of frozen blocks of beef using the percent fat as the only variable. When fat in beef has to be closely controlled in the finished product a sampling plan for incoming lots of beef is required. Sampling techniques, chemical procedures and charts showing the admissible errors within certain confidence limits for different sampling rates will be discussed. The uses and advantages of such a plan will be discussed.

ACKNOWLEDGMENT

The author acknowledges the guidance of Dr. Gordon A. Rowe, University of California, Agricultural Extension Service in the theoretical formulations of the problem. Co-workers of the Gerber Quality Control staff were Philip Johnston, Gordon Everett and Darrell Brown.

Frozen beef is purchased in the food industry for many purposes, particularly as raw products in the manufacture of formulated items. As in all other raw materials, certain quality factors must be maintained and controlled and this is best accomplished by manufacturing or in this case "boning" them in rather than inspecting them in. In the case of beef when used in hamburger, frankfurters and similar commodities, it becomes important to know the fat content of the lots purchased. It is likewise important to know the protein and moisture contents plus the sanitary factors. For purposes of this paper, the fat content will be considered as the only variable.

In our case, maximum fat content for raw beef is specified to meet the rigid requirements of our finished products. The problem then is to estimate the fat content of each incoming lot by sampling. Beef is usually received in 100# boxes and in 200 box lots sizes. The meat is sampled by drilling into the frozen block with a drill and receiving the drilling, mixing it by hand and analyzing a 9 gram sample using the Babcock fat test.

The fat in each block and between blocks is highly variable, varying in degree from ground beef to much more variable cuts such as the beef chux. As a contrast, milk is usually tested for fat using the Babcock test. However, milk is sampled from a population that is relatively homogeneous because of its physical character and the constant mixing occurring in transit or in temporary storage. The inherent variability of fat in a beef carcass may be to a larger extent maintained in the boned lots. In estimating the true value of such lots, variability as indicated above becomes an important aspect of the problem of achieving the desired precision of estimation.

Inasmuch as fat varies between cores in a block and between blocks, it may be necessary to estimate the fat content by drilling one or more cores for several blocks. This then could be the basis for subsequent acceptance of rejection of the lot. We must also keep in mind that we are interested only in the upper limit of fat content. We are interested in a one sided test. We are then interested in the following formula:-

$$p_{\bar{x}} \leq P_{\bar{x}} + k$$

when $p_{\bar{x}}$ = estimate of the true fat content, $P_{\bar{x}}$, of the lot and k is an admissible sampling error.

TABLE I
Analysis of Variance

Core	Block A	B	C	D	E	F	G	H	I	J	Total
1	10.1	10.4	12.5	10.4	12.0	12.0	10.6	13.8	10.5	9.5	
2	17.0	16.0	12.5	10.6	10.2	14.0	10.8	12.8	12.0	10.0	
3	13.0	13.0	12.0	12.0	9.6	11.0	9.0	11.8	11.0	11.5	
4	10.5	12.8	12.5	11.4	10.0	10.0	11.0	10.2	11.2	9.5	
5	15.5	10.6	11.5	9.4	11.6	14.4	12.0	14.2	12.0	10.5	
6	12.5	12.2	10.5	10.2	13.0	13.6	11.0	12.0	13.4	10.5	
$\sum X$	.786	.750	.715	.640	.664	.750	.644	.748	.701	.615	7.013
$\bar{X}$	.131	.125	.119	.107	.111	.125	.107	.125	.117	.103	
$\sum X^2$	.106676	.095820	.085525	.068688	.074376	.095332	.069600	.094320	.082425	.063325	.836087
$\frac{(\sum X)^2}{n}$	.102966	.093750	.085204	.068267	.073483	.093750	.069123	.093251	.081900	.063038	.824732

	Sums of squares	df	mean square
Means, between	.005029	9	s_b^2 .000559
Within	.011355	50	s_w^2 .000227
Total	.016384	59	

$$s_u^2 = \frac{s_b^2 - s_w^2}{m}$$

$$= .000055$$

$$s_w^2 = .000227$$

We assume normal distribution of the sample means. The variance of this distribution, $s^2_{\bar{x}}$, is:-

$$\frac{N-n}{N}\ \frac{s_u^2}{n} + \frac{MN-mn}{MN}\ \frac{s_w^2}{mn}$$

where N = blocks in a lot (200)
M = cores in a block (100)
n = blocks to be selected in a test
m = cores to be sampled in each block
s_u^2 = estimate of the sampling distribution between block mean.
s_w^2 = estimate of the sampling distribution within blocks.

The estimating formula then is:-

$$k^2 \leq z^2 \left(\frac{N-n}{N}\ \frac{s_u^2}{n} + \frac{MN-mn}{MN}\ \frac{s_w^2}{mn}\right)$$

where k is the admissible in estimating fat content and z is our confidence limits. For simplification of the problem corrective factors for finite population on the left side of our two components of variants may be omitted. Our estimating formula then is:-

$$k^2 = z^2\left(\frac{s_u^2}{n} + \frac{s_w^2}{mn}\right)$$

We assume the variance components are known. We can solve for n and m given k and z and estimate of s_u^2 and s_w^2. Solution for n and m are as follows:-

$$n = \frac{s_u^2 + \frac{s_w^2}{m}}{\frac{k^2}{z^2}}, \qquad \text{m given and}$$

$$m = \frac{s_w^2}{n\frac{k^2}{z^2} - s_u^2} \qquad \text{n given}$$

Various estimates:- s_u^2 and s_w^2 estimates are made by analysis of variance techniques. Ten blocks are drilled 6 times each and the individual drillings are analyzed for fat content. Table I gives the results from a lot of ground beef. It is recognized that samples drawn from other lots will vary from this, assuming variations similar to those in Table I, we can use these to estimate our variance and set up sampling plans for lots of this type.

From Table 2 a sampling plan can easily be chosen with admissible errors of 1%, 2% and 5%. For example, two blocks each drilled once would give a result with an admissible error of 2%. To get an admissible error of 1% we would have to drill 7 blocks with 2 cores each or 8 blocks of one core each.

It is to be recognized that the above is on one lot only. Further investigation may reveal different variations depending on possible boners, type of cattle, the mixing qualities of the mixer and other factors.

Sampling Plans

The sampling plan can be selected according to the desire of the one administrating the sampling and dependent on how precise a result be needed at what confidence limits he has to work to. We would want to know that with a maximum specification of 12% fat what percent of the time a 11% fat content would be inside the limits for the true value. In general we can say using k = (.01) that if the fat value is 10% then 95% of the time the true fat value

is 10%, then 95% of the time the true fat value of the lot will not exceed 11%. The optimum sampling plan then gives us the minimum number of blocks and cores to sample with varying degree of k_u^2 .

TABLE II
Alternative Sampling Plans

blocks n	k	z = 1.645 (95%) .01	.02	.05	cores m	z = 1.960 (99%) .01	.02	.05
1		0	0	1		0	0	1
2		0	1	1		0	0	1
3		0	1	1		0	1	1
4		0	1	1		0	1	1
5		0	1	1		0	1	1
6		0	1	1		0	1	1
7		2	1	1		0	1	1
8		1	1	1		0	1	1
9		1	1	1		8	1	1
10		1	1	1		2	1	1

Prior to the analysis a sampling plan was used by drilling six cores from each of 8 blocks. By using the second formula and z = 1.645 we find k to be 0.57%. This is to be compared with 7 blocks drilled twice each to give a k of 1% at 95% confidence limits. It is obvious that we were over sampling.

$$k^2 = 1.645^2 \left[\frac{.000055}{8} + \frac{.000227}{(8)(6)} \right]$$

$$k = .0057 \quad (95\% \text{ confidence limits})$$

There are many advantages and uses to which this can be put to if one is engaged in the buying of frozen beef:-

1. Accept or reject on the basis of fat content with known limits of confidence.
2. Use these results in blending to attain desired fat content in final results.
3. Set up a premium and discount system for various values of fat content.
4. Afford an optimum sampling to get desired results.
5. Gives one a real basis on which to compare suppliers, source of supply, cuts of meat and in general save dollars in the purchase of a very valuable item in the formula for food processing.

SELECTION OF FLAVOR PANELS FOR COMPLEX FLAVOR DIFFERENCE

Mae-Goodwin Tarver and Barbara Hall Ellis
Metal Division Research & Development Department
Continental Can Company, Inc., Chicago 20, Illinois

INTRODUCTION

Physical instruments are not available at this time to detect small differences in flavor of a food product. Human response, therefore, is still the most sensitive measure of flavor quality. Response to a flavor difference depends upon: (1) the precision or the inherent ability of a taster to duplicate a difference judgment; (2) the reliability of a taster or the absence of bias when detecting a flavor difference; and (3) the tolerance level of a taster or his inherent sensitivity to particular flavor differences. These three factors must be considered when flavor panel members are selected for flavor difference tests. They must also be considered when the data resulting from these difference tests are submitted to statistical analysis.

In two previous publications [1,2] a method was described for selecting beer panel members for difference tests based upon their sensitivity to bitterness. A method of this type is efficient and practical when the predominant flavor of the food or beverage is easily characterized by one of the four basic tastes. The flavor of many food products, however, is due to a complex mixture of odors, tastes, and mouth-feelings. No one basic taste is characteristic of this complex mixture. The purpose of this paper is to describe a method for selecting flavor panel members for difference tests when complex flavors are present. Another purpose is to illustrate the statistical analysis of data obtained from these flavor difference tests. Data resulting from Two-Sample Difference Tests upon experimental packaging materials are used to illustrate the methods.

FOOD PRODUCT AND TEST METHOD

Frozen orange juice concentrate was selected for this study of experimental packaging materials because the characteristic flavor of this product contains sweet, sour, and aromatic components. The frozen concentrate was reconstituted to its original strength before serving. The flavor test method used in this study was the Two-Sample Difference Test.[2]

EXPERIMENTAL DESIGN

Three experimental packaging materials, i.e., experimental containers, and a reference standard, i.e. the test control or commercial container, were packed in contact with frozen orange juice concentrate.

A pair of samples, the control and a test variate, were served to each of 13 available tasters. In addition, a control-to-control pair were tasted by each of the tasters. Each taster was asked to determine the degree of difference between each pair of samples using the following difference scale:

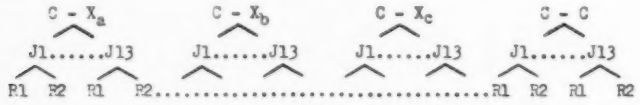
Score	Description
0	No difference detected
1	Very slight difference
2	Slight difference
3	Moderate difference
4	Large difference
5	Extremely large difference

On the data sheets, this degree of difference scale was printed as a vertical structured scale.[2] This type of scale arrangement encourages the use of fractional difference scores, e.g., 0.25 or 0.50. Tasters were permitted to use fractional difference scores because previous studies indicated that such a response to flavor differences is characteristic of an individual taster and is reflected in his precision.[1]

Duplicate sets of each pair were included in the test. However, only two pairs (four samples) were tasted at any one time to avoid taste fatigue. The duplicate pairs of a specific test variate, therefore, were randomized with respect to time of tasting.

The control sample of a pair was identified for each taster. This is a desirable procedure since this reference point aids the taster in making accurate flavor-difference judgments. Taster bias, which can be created by this prior knowledge of the test, was evaluated by the control-to-control comparisons included in the test.

The formal experimental design of this Two-Sample Difference Test was the following nested (hierarchal) design:



Legend

J = Taster number = 13
X_a, X_b, X_c = Experimental packaging materials = 3
C = Reference or control sample
R1, R2 = Replications = 2

The special features of this design permitted (1) the detection of flavor-difference response in the precision of the tasters (mean square of the duplicate samples) and (2) an appraisal of the flavor-difference characteristics of the experimental packaging materials by regression analysis if the tasters showed a significantly different response within the materials.

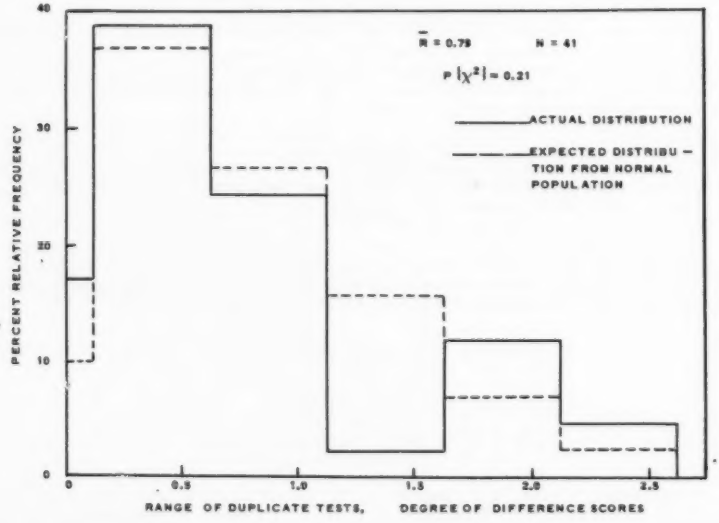


Figure 1. Distribution of ranges of duplicate tests: Two-Sample Difference Flavor Test. (13 Tasters - Three packaging materials plus reference sample.)Orange juice concentrate

EVALUATION OF TASTER PRECISION FROM CONTROL-TO-CONTROL COMPARISONS

Experimental evidence accumulated in this laboratory shows that the response to a flavor-difference sometimes appears in the precision of the tasters' difference scores as well as in the average difference scores.[1] For this reason, the inherent precision of the flavor panel must be evaluated in the absence of flavor-difference response, i.e., from the control-to-control comparisons included in the tests.

Statistical appraisal of taster precision assumes that the data were drawn from a normal population. Previous experimental data resulting from the Two-Sample Difference Test have shown no evidence of non-normality of the flavor-difference scores.[1] In this specific test, again no evidence of non-normality was observed. Figure 1 compares the actual distribution of the ranges of the duplicate samples run in this test with the expected distribution of the ranges if the samples were drawn from a normal population.[3] A Chi-square goodness-of-fit test between the actual and the expected distribution resulted in a probability of 0.21. This indicated no significant departure from normality of the flavor-difference scores. Therefore, the precision of the 13 tasters was evaluated by a statistical control chart for the range of the control-to-control pairs. This range chart, shown in Figure 2(a), indicates that three tasters (Nos. 1,9, and 10) were not as precise as the remaining 10 tasters. The ranges of these three tasters lay above the upper limit of the chart. The flavor-difference scores obtained from these three tasters were, therefore, deleted from the data in subsequent computations.

APPRAISAL OF TASTER BIAS

The control chart for averages shown in Figure 2(b) assumes that the average difference resulting from a control-to-control comparison is expected to be zero if the taster is completely unbiased. However, chance fluctuations in the taster's judgment introduce a source of random bias. The upper limit for an average difference of zero is, therefore:

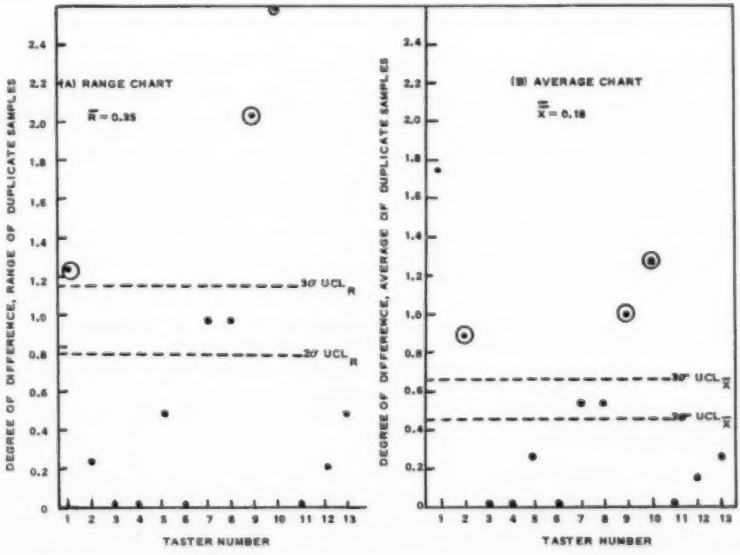


Figure 2. Statistical Control Charts for Control-to-Control Comparisons. Selection of a precise, unbiased panel. (Orange juice concentrate)

$$0 + A_2\bar{R} = 0 + 3\frac{\bar{R}}{d_2\sqrt{n}}$$

Where: A_2 = a factor related to sample size (4)
$\bar{R}$ = average range
d_2 = a factor related to sample size (4)
n = sample size
3 = a multiplier related to 99.73% probability

If an average difference score for a control-to-control pair lies above this upper limit, a taster is very likely to be significantly biased or to be affected by factors other than those included in the test conditions. Four biased tasters (No. 1, 2, 9, and 10) are shown in Figure 2(b) and the flavor-difference data of the tasters were eliminated in subsequent statistical analyses of the data obtained from the experimental packaging materials.

The control-to-control difference scores of the remaining nine panel members were submitted to the analysis of variance. The results of this analysis are shown in Table I and they indicate that a homogeneous panel of nine tasters was obtained in this test using orange juice concentrate.

TABLE I
ANALYSIS OF VARIANCE OF FLAVOR DIFFERENCE DATA. CONTROL-TO-CONTROL COMPARISONS. ORANGE JUICE CONCENTRATE

SOURCE OF VARIATION	DEGREES OF FREEDOM	MEAN SQUARE	F-RATIO
AMONG TASTERS	8	0.0868	0.61 NS
WITHIN TASTERS (RANDOM ERROR OF DUPLICATES)	9	0.1424	--

PANEL AVERAGE = 0.18; NS = NOT SIGNIFICANT (BELOW 5% LEVEL)

TASTERS' TOLERANCE LEVELS

If there is a significant difference among the difference scores assigned to one experimental packaging material, then the initially homogeneous panel of tasters is responding differently to the applied flavor-difference stimulus. The conclusions drawn from the statistical analysis shown in Table II are: (1) that the nine tasters have significantly different tolerance levels for the packaging materials and (2) that some of the flavor-difference response is appearing in the precision of the tasters. The first conclusion means that the tolerance level of each taster for each different packaging material must be calculated, while the second conclusion means that the random error for statis-

TABLE II
ANALYSIS OF VARIANCE OF TWO SAMPLE DIFFERENCE TEST DATA. THREE EXPERIMENTAL PACKAGING MATERIALS. ORANGE JUICE CONCENTRATE

SOURCE OF VARIATION	DEGREES OF FREEDOM	MEAN SQUARE	F-RATIO
AMONG PACKAGING MATERIALS	2	0.1830	--
AMONG TASTERS WITHIN MATERIALS	24	0.1830	5.04**
REPLICATIONS WITHIN TASTERS	27	0.4410	3.10*
DUPLICATE ERROR FROM CONTROL-TO-CONTROL	9	0.1424	--

* = 5% LEVEL (TREND); ** = LEVEL (CRITICAL LEVEL)

tical tests of significance must be obtained from the control-to-control comparisons shown in Table I. The tolerance level of a taster to a specific packaging material is defined as the taster's ability to detect a difference between the control sample and a sample of the packaging material under test. The statistical F-ratio was satisfactorily used by R. G. H. Prince and J. H. Ince[5] to measure sensitivity to odors and was also used in this flavor-difference study to reflect the tolerance of each taster to the flavor-difference of the packaging materials. The F-ratio was calculated by the following equations:

$$F = \frac{MS_B}{MS_E} \qquad (2)$$

Where: MS_B = mean square between the control and the packaging material

MS_E = error mean square

$$MS_B = \frac{(\sum x_{c-c} - \sum x_{c-x})^2}{2n} \qquad (3)$$

Where: $\sum x_{c-c}$ = sum of control-to-control comparisons

$\sum x_{c-x}$ = sum of control-to-variate comparisons

n = number of replications of each comparison

$$MS_E = \frac{\sum R^2}{2k} \qquad (4)$$

Where: $\sum R^2$ = sum of (ranges)^2 of control-to-control and control-to variate duplicate samples

(note: for duplicate samples only) k = number of ranges summed

The F-ratios calculated in this manner are shown in Table III for each taster and for each of the three experimental packaging materials. Individual tasters do not necessarily have the same tolerance level for the three different packaging materials.

TABLE III

DATA FROM TWO-SAMPLE DIFFERENCE TEST USING HOMOGENEOUS, UNBIASED PANEL OF NINE TASTERS. ORANGE JUICE CONCENTRATE

TASTER NUMBER	EXP. MATERIAL "A"		EXP. MATERIAL "B"		EXP. MATERIAL "C"	
	F-RATIO	*DIF. SCORE	F-RATIO	*DIF. SCORE	F-RATIO	*DIF. SCORE
3	0.05	0	0.98	1.75	0.28	0.75
4	0.05	0	0.50	0.25	0.05	0
5	0.74	1.500	0.50	0.50	0.25	0.50
6	0.50	0.250	0.05	0	0.50	0.50
7	0.90	1.250	0.12	0.25	1.12	1.25
8	0.50	0	0.50	1.00	0	0.50
11	1.00	2.000	1.00	1.50	1.28	0.75
12	0.50	0.375	0.10	0.25	0.10	0.25
13	0	0.250	0.25	0.50	0	0.25

*AVERAGE OF DUPLICATE TWO-SAMPLE DIFFERENCE TESTS

STATISTICAL INTERPRETATION OF THE TEST DATA

The average flavor-difference scores assigned to the three experimental packaging materials are also shown in Table III. Since the analysis of variance shown in Table II indicates that a significant difference among tasters exists, differences among the three experimental packaging materials were determined by regression analysis. In the three separate regression analyses given in Table IV, the differences among tasters were partitioned into (1) the regression due to differences in tolerance levels and (2) the unexplained variation or the residual variance about the regression lines. Figure 3 illustrates the regression of the average difference scores of the three different packaging materials against the taster's tolerance levels. Each packaging material shows

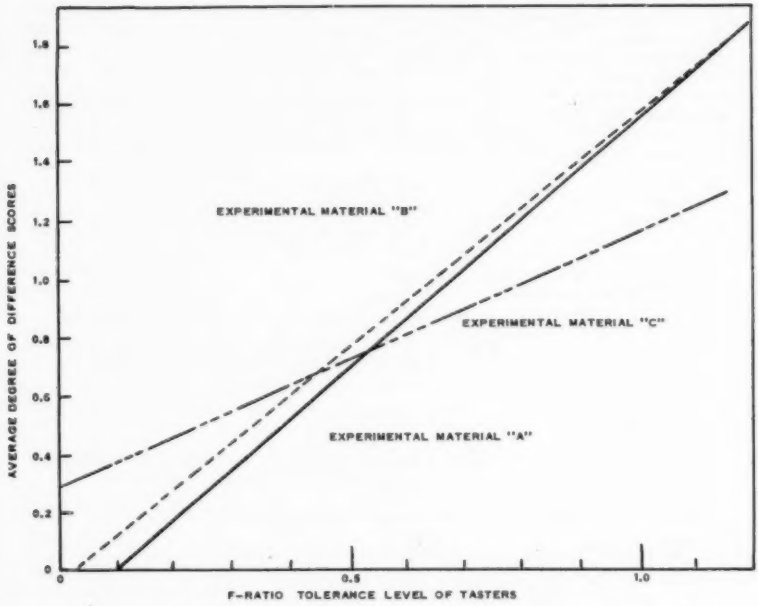


Figure 3. Individual regression lines for Two-Sample Difference Flavor Test. Three experimental packaging materials. Orange juice concentrate.

TABLE IV
ANALYSIS OF VARIANCE OF INDIVIDUAL REGRESSIONS, F-RATIOS AND DIFFERENCE SCORES OF EXPERIMENTAL PACKAGING MATERIALS, ORANGE JUICE CONCENTRATE

	EXP. MATERIAL "A"			EXP. MATERIAL "B"			EXP. MATERIAL "C"		
SOURCE OF VARIATION	DEGREES OF FREEDOM	MEAN SQUARE	F-RATIO	DEGREES OF FREEDOM	MEAN SQUARE	F-RATIO	DEGREES OF FREEDOM	MEAN SQUARE	F-RATIO
AMONG TASTERS	8	1.1406	–	8	0.7500	–	8	0.2648	–
DUE TO REG.	1	6.4876	16.98**	1	5.5338	80.09**	1	1.4784	16.35**
DEV. FROM REG.	7	0.3810	3.68NS	7	0.0686	0.47NS	7	0.0904	0.63NS
DUP. ERROR FROM CONTROL-TO-CONTROL	9	0.1424	–	9	0.1424	–	9	0.1424	–
(A) PANEL AVERAGE		0.82			0.67			0.53	
COMPONENTS OF FLAVOR CONTOURS	F-RATIO	(1) BASIC TASTER	(2) MISC. FACTORS	F-RATIO	(1) BASIC TASTER	(2) MISC. FACTORS	F-RATIO	(1) BASIC TASTER	(2) MISC. FACTORS
(B) % OF TOTAL MEAN SQUARES	94.42%	2.09%	3.49%	94.04%	5.96%	0%	91.21%	8.79%	0%
(C) (A) X (B) COMPONENT OF AVG. SCORE	0.59	0.01	0.02	0.63	0.04	0	0.48	0.05	0

NS=NOT SIGNIFICANT (BELOW 5% LEVEL): * = 5% LEVEL (TREND): ** = 1% LEVEL (CRITICAL LEVEL)
(1) BASIC TASTER = MS OF CONTROL-TO-CONTROL: TRUE RANDOM ERROR OF TASTERS FROM TABLE I
(2) CALCULATED BY ASSUMING MS OF CONTROL-TO-CONTROL CONTAINS ALL RANDOM ERROR OF TASTERS:
$MS_{DEV.} - MS_{DUP.}$ = EXCESS MS = MISC. FACTORS OR FACTORS NOT ACCOUNTED FOR

TABLE V
ANALYSIS OF COVARIANCE. TWO-SAMPLE DIFFERENCE TEST OF THREE EXPERIMENTAL PACKAGING MATERIALS. ORANGE JUICE CONCENTRATE

SOURCE OF VARIATION	DEGREES OF FREEDOM	MEAN SQUARE	F-RATIO
IMPROVEMENT IN ANALYSIS DUE TO USE OF REGRESSION (OVER TABLE II)	1	12.1278	54.61**
DEVIATIONS FROM REGRESSION	23	0.2221	1.86NS
DIFFERENCE AMONG SLOPES OF INDIVIDUAL REGRESSION	2	0.7209	4.02*
"A" VS. "B"	1	0.0564	0.32NS
"A" + "B" VS. "C"	1	1.3884	7.72*(1)
DEVIATION FROM INDIVIDUAL REGRESSION	21	0.1794	1.25NS
DUPLICATE ERROR FROM CONTROL-TO-CONTROL	9	0.1424	--

NS = NOT SIGNIFICANT (BELOW 5% LEVEL); * = 5% LEVEL (TREND); ** = 1% LEVEL (CRITICAL LEVEL)
(1) F-RATIO IS VERY CLOSE TO THE CRITICAL LEVEL

a significant correlation of difference scores with taster tolerance levels and the slope of each line is significantly different from zero, i.e., a significant difference was detected between the control sample (reference standard) and the three experimental materials. The true random or error term used in each analysis of variance was obtained from Table I.

An over-all comparison of the three different materials was made by the analysis of covariance (6) shown in Table V. There was no evidence of a significant difference between the slopes of Materials "A" and "B", indicating that the tasters reacted in a similar manner to these two materials. However, these two slopes were significantly steeper than the slope of the regression line for Material "C". The experimental interpretation of this difference among slopes was that the panel members found significantly larger flavor-differences, per

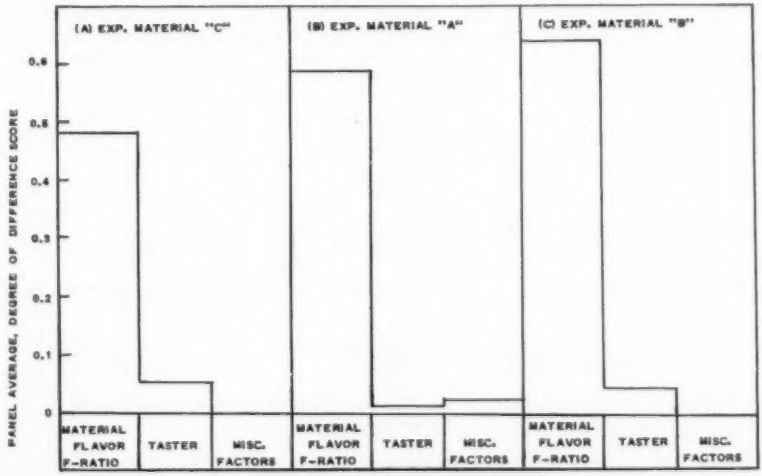


Figure 4. Continental flavor-difference contours for three experimental packaging materials. Orange juice concentrate

unit tolerance level, between Materials "A" and "B" and the control or reference standard than between Material "C" and the control. Since Material "C" was the nearest approximation to the reference standard, it was retained for further testing. Materials "A" and "B" were deleted from the flavor-difference tests since additional information contained in cuality scores indicated that the observed difference from the reference standard was a negative difference.

CONTINENTAL FLAVOR-DIFFERENCE CONTOURS

The flavor-difference contours, shown in Figure 4, were calculated from the information given in Table IV. The contours for Materials "A" and "B" show similar characteristics but are different from the contour for Material "C". This latter contour has a somewhat shorter Material Flavor bar, indicating that it is a closer approximation to the control or reference standard.

SUMMARY

The precision and the bias of flavor panel members can be objectively appraised by standard statistical methods even in the presence of complex flavors. Unbiased tasters of similar precision can be selected under these conditions using machinery or instrument capability study technicues.

Tolerance levels for complex flavor-differences can be calculated for individual tasters for each material or product being tested. The F-ratio, based on control-variate comparisons, is a satisfactory measure of this tolerance level.

Differences among test variates can be evaluated by regression and covariance techniques. These methods are based upon the association between taster tolerance levels (F-ratios) and flavor-difference scores. Differences among taster tolerance levels, therefore, do not interfere with the interpretation of complex flavor-differences, provided that the tasters initially constitute a homogeneous, unbiased panel.

Flavor-difference contours can be drawn to show the principal components of complex flavor-differences. These contours are based upon the regression analyses employed in the appraisal of the data and permit visual comperisons of flavor-differences among the test materials. The contours are similar to those described in previous publications (1),(2) although the contours described in this current paper are limited to a maximum of three components.

ACKNOWLEDGEMENTS

The authors wish to thank Dr. N. E. Henry and Dr. B. H. Morgan for their assistance in the preparation of this paper. They also wish to thank Mrs. A.M. Schenck for illustrations contained in this paper.

REFERENCES

1. Tarver, Mae-Goodwin, Hall, Barbara A., and McDonald, Joan G. A Statistical quality control approach to the selection of flavor panel members. Transactions, 13th Annual Convention, ASQC, pp. 459-485, (1959).

2. Hall, Barbara A., Tarver, Mae-Goodwin, and McDonald, Joan G. A method for screening flavor panel members and its application to a two-sample difference test. Food Technol. 13:12, 699-703, (Dec. 1959).

3. Private communication. Dr. W. J. Youden to Dr. Marjorie L. Sutherland, (1954).

4. Grant, E. L. Statistical Quality Control (2nd edition). New York: McGraw-Hill Book Company, pp. 512-513, (1952).

5. Prince, R. G. H. and Ince, J. H. The measurement of the intensity of odour. J. Applied Chem. 8:5, 314-321, (1958)

6. Snedecor, George. Statistical Methods (5th edition). Ames, Iowa: Iowa State College Press, pp. 395-404, (1956).

HOW MANY DOLLARS ARE THERE IN INCOMING INSPECTION COSTS?

F. Ennerson, Supervisor, Statistical Analysis
Hughes Aircraft Company, Tucson, Arizona

Purchasing must have immediate and dependable quality and reliability data upon which to base a procurement decision.

Statistical techniques are required to evaluate the test results of purchased parts in order to advise Purchasing in procurement of future materials.

In many manufacturing operations, operating and procurement costs will reveal that material costs are 60% to 70% of the total cost of the manufacturing operation with the actual manufacturing operation costs between 10% and 15%.

"How many dollars are there in incoming material costs?"-- the most fertile area for cost saving is in the incoming material.

1. Dollars saved on streamlining paperwork in Receiving Inspection. Receiving Inspectors should check the product instead of spending 25% to 50% of their time bookkeeping.

2. Dollars saved on determining in advance what material is required to satisfy design requirements.

3. Dollars saved ordering in economical lot sizes.

4. Dollars saved in properly written instructions for the inspector which state what tests and what equipment should be used and what classification of defects shall apply.

5. Dollars saved in having properly trained inspection personnel.

6. Proper sampling plans with acceptable producer and consumer risks.

7. Dollars saved using process average information for each vendor part number for statistically sound selection of reduced, normal, and tightened sampling plans.

8. Dollars saved by processing inspection results in order that accounting may take advantage of discounts for early payment. Establish a system for sampling cost accounting paperwork. Stratify by cost of item and sample all but high-cost-item paperwork.

9. Dollars saved in having realistic specifications.

10. Dollars saved in having approved vendors with acceptable Quality Control operations.

11. Dollars saved in having vendors send in their inspection results with the material.

12. Dollars saved in having environmental test results prior to acceptance of material.

13. Dollars saved in monitoring factory and field failures for feed-back information to vendor and purchasing.

14. Dollars saved in having an effective vendor rating system.

HOW MANY DOLLARS ARE THERE IN INCOMING INSPECTION COSTS?

To get the benefit of these dollars does not in most cases require initiation of a new system of data collection. Information on incoming lots is generally readily available if for no other reason than to determine whether or not to pay the vendor for material received. It is in scientifically using the data that costs savings ensue.

There is a well-known axiom that "Quality Products" will cost you more. The better the material, the finer the workmanship, the more thorough the service -- these things all increase the price. However, keep in mind that the quality of an item refers to its adaptation to a particular need.

A designer may stress durability in a product, or long life expectancy, or simplicity, or maintainability, or a combination of such features. Each of these features affects cost in one way or another. Now it is inspection's prime function to ascertain product conformance to the design.

To achieve continual conformance to design objectives, there must be a high degree of cost consciousness in the efforts of both the purchasing and inspection departments. Purchasing obviously wants to buy at the lowest possible unit cost but must also consider product availability and quality. Inspection, of course, wants acceptable product delivered to the receiving dock to take advantage of cost savings in the form of reduced inspection.

Some advantages of close inspection-purchasing coordination might be in permitting larger quantities to be bought on the basis of a good unit history with resultant savings through increased discount. Similarly, statistical reviews such as process capability study may indicate selecting a wider tolerance, less expensive purchased part to be used in an assembly. Further, reliability analysis may indicate the need for better material content than the designer originally specified and which purchasing subsequently obtained.

Buyers, of course, have access to cost and product availability data. However, inspection data may not be known unless a system is devised to furnish it to the buyers when placing an order. In the majority of instances buyers are not told exactly what brand of product to buy from whom. Usually, a request is made to the purchasing department to obtain 280# of SAE 1020 Steel, or 500 10,000 ohm resistors. Consequently, the buyer's ability to obtain a satisfactory product from a reliable vendor is often based on cost and availability factors alone.

The need for some sort of firm tie-in is evident; both before the material is ordered, after it has been received and inspected, and after subsequent material usage in the manufacturing process.

During the original placing of an order, it is important that the required quality be adequately specified on the purchase document. Complete definitive terms are needed to make sure that the vendor knows exactly which quality characteristics are expected. It is also important that prior knowledge of what is being purchased be made available to the inspection department to be sure that inspection can be performed. If a particular order is going to require special gaging to determine its acceptability, the inspection department should take steps to obtain the required gages before the material arrives. Such prior knowledge also enables the detailing of the inspection procedure to be used so that a reasonable amount of inspection is performed. Too much inspection increases labor costs and affects the scheduled release of material. Too little inspection can lead to possible defective parts in production and subsequent bottlenecks.

HOW MANY DOLLARS ARE THERE IN INCOMING INSPECTION COSTS?

PROCESS AVERAGES AND VENDOR RATING

Ideally, of course, all material would be brought into the receiving dock with enough confidence in its quality to send it directly to stores. This millenium has not arrived, however, and receiving inspection is still necessary. Nevertheless, the quality department can take several forward steps in the direction of the millenium in the form of vendor qualification surveys to see that the vendor has the organization, the equipment, and the necessary procedures for controlling product quality. In many cases the vendor is being required to certify that the material he furnishes will meet specifications and prove that he has performed the necessary inspection and test to give this assurance.

Let us admit that most of us are still performing a large amount of receiving inspection. Most of us are performing receiving inspection on a sampling basis because we haven't time or money to inspect every screw and washer 100%, and we know that hundred-percent inspection is only eighty percent efficient anyhow. However, far too often our sampling inspections are not being performed scientifically. Sampling involves a calculated risk, and you've got to understand exactly what the risks are and how to minimize them. Here is where statistical analysis can really pay off in dollar savings by answering two basic questions:

1. What is the minimum amount of testing and inspection that will assure the required quality level, meanwhile preventing lots of marginal quality from getting into stores?

2. Which vendors can be most depended upon to furnish material requiring the least amount of test and inspection?

To answer the question about the minimum amount of testing and inspection, we can first classify the relative importance of possible defects and assign AQL's for each class. Then we choose scientific sampling plans and apply them properly.

The classification of defects (C/D) is often made by receiving inspection planning engineers according to such established criteria as MIL-STD-105 which defines Critical, Major, and Minor defect classes. Or you might devise a system of your own, similar to the following:

Class 0

Will result in a hazardous or unsafe condition for personnel during handling, installation or end use.

Class 1

a. Will prevent the system from performance of its function and is not likely to be found at a subsequent functional test prior to delivery.

b. Will cause intermittent operating trouble difficult to locate in the field.

c. Will fail to meet shelf-life requirements.

d. Will interfere with required installation for end use or prevent interchangeability where required.

Class 2

a. Not likely to result in functional failure but will result in materially reduced efficiency of operation and is not likely to be found in a subsequent functional test prior to delivery.

b. Will cause intermittent operating trouble which can be readily corrected in the field.

c. Will result in considerably decreased life of the product.

d. Will result in excessive maintenance.

e. Will cause minor installation difficulties or adversely affect interchangeability where required.

f. Will be found in subsequent test; is excessively time-consuming or costly to correct.

Class 3

a. Likely to cause trouble of a nature less serious than an operating failure.

b. Important defects of appearance, finish, or workmanship.

c. Will result in moderately increased maintenance or decreased life representing some change from normal expectancy.

d. Variations from engineering requirements in detail parts or assemblies that affect only ease of assembly or minor adjustment.

e. Will be found in subsequent test; is not time-consuming or costly to correct.

Class 4

a. Will have no significant effect on operation, functional use, or life of product in service.

b. Will not interfere with installation or adversely affect assembly operations.

c. Minor defects of appearance, finish, or workmanship.

Acceptable Quality Levels (AQL's) assigned for such C/D classes might be as follows:

C/D Class	AQL
0	0% (100% inspection)
1	0.4%
2	1.0%
3	2.5%
4	6.5%

Once the C/D classes and Acceptable Quality Levels have been assigned, scientific sampling plans can be chosen such as Dodge-Romig, MIL-STD-105, or various refinements of these. Once chosen, these sampling plans should be

applied properly. By "applying them properly" I mean keeping process average histories and employing properly trained inspection personnel. An inspector should have a knowledge of how sampling plans operate and the importance of following the sampling plan. This requires at least a nodding acquaintance with basic statistical principles. (Be sure that your inspection department foreman knows something about statistics, too).

Let us emphasize that the true quality level of a single lot of material inspected on a sampling basis is still nothing more than a rough guess. Only by the use of the process average history over a series of successive lots can we establish any high degree of confidence in the measure of the true quality level. Yet there are still organizations that do not keep process averages on incoming parts.

When process average histories are in use, as many as four different sampling inspection levels may be used with one sampling plan, Tightened, Normal, Reduced and Double Reduced. Where these histories are kept, it becomes statistically sound and a definite saving to apply reduced sampling on parts that have good process averages. It may even be possible to reduce sampling inspection to one in every five lots of incoming material with the intervening lots inspected only for identity, count, and obvious damage. Some of the time saved by not performing unnecessary inspection on good parts can be utilized in giving suspect or marginal material a more thorough inspection on a Tightened sampling level. Depending on the OC Curve, the probability of accepting defective material is altogether too high in many instances. In order to reduce this probability of acceptance, it is quite necessary to use Tightened sampling.

Is it difficult to keep process average histories? No, keeping a process average history sheet on inspected lots of the same part number from the same vendor is essentially a simple operation. The mechanics of it are merely to note the size of the shipment or lot presented for inspection, record the number inspected in the first sample for each C/D class, the number found defective in that sample, and whether the lot was suspended or accepted. The results of the second sample where double sampling is used are disregarded except for the decision to accept or suspend the lot. Inspection results are tabulated for each C/D characteristic (mechanical, electrical or visual) and class. It is possible that records may be kept for as many as nine (9) characteristics on a particular vendor part.

Determination of future inspection levels is made following each lot received and inspected by reference to the recorded process average history and a process average table. It is possible to go to or retain reduced sampling inspection levels with an occasional defective as limited by the particular process average table. However, any lot suspension should automatically return the sampling level to Normal. A number of defectives greater than the upper limit for Normal sampling results in the assignment of Tightened sampling for the next lot received.

As stated previously, it is not difficult to keep process averages. Once the system has been set up with a few simple tables, a clerk can perform the entire operation. In a small or medium-sized organization, manual operation will be adequate. It may not even be necessary to use an adding machine. One employee full time can handle an active file of several thousand parts where the inspection load does not exceed five to seven-hundred (500-700) lots per week. Beyond this volume, the use of automatic tabulating equipment is indicated.

HOW MANY DOLLARS ARE THERE IN INCOMING INSPECTION COSTS?

Now to answer Question Two: "Which vendors can be most depended upon to furnish material requiring the least amount of test and inspection"? Vendors can be rated directly from the process average history sheets at whatever rating interval may be desired -- even lot-by-lot. They can be rated by such a system as:

A - Preferred
B - Good
C - Acceptable
S - Suspended
X - Disapproved

These ratings can be assigned in accordance with sampling levels in effect on each vendor part. Weighting factors are deemed unnecessary because we expect the vendor to live up to all the respective AQL's assigned. A string of "Tightened" inspection levels on any characteristic should result in the suspension or disapproval of a vendor. (Of course your Quality Department should make every effort to assist the vendor in producing good material, since it is advantageous to get and keep good vendors.) On the other hand, assignment of "Double-Reduced" or "Skip-Lot" inspection levels should earn a "Preferred" rating.

The Purchasing Department should receive copies of all rating changes and use them when ordering. An agreement to require special negotiation with the quality department to place an order with a "Suspended" or "Disapproved" rated vendor can be worked out.

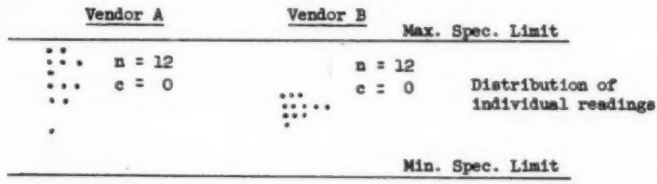
No vendor rating system can be effective unless it is up-to-date and the purchasing department respects its validity.

ENVIRONMENTAL TEST DATA

Environmental tests are required by many part specifications. Usually these tests are conducted by laboratories which are apart from the receiving inspection area. Current environmental requirements are on an attribute basis, and in addition require relatively small sample sizes of six (6) through twenty-four (24). Since environmental tests are expensive (due to manhours, equipment requirements, and length of time) more effort is required in getting maximum information out of the data. The days of go, no-go data are limited, as we do not know too much more about the lot after the tests are completed. Variables data, of course, is not novel, but it is a must in environmental test data. It is necessary to know not only if the sample meets the specification requirement but how good or how bad the sample is in reference to the specification requirements.

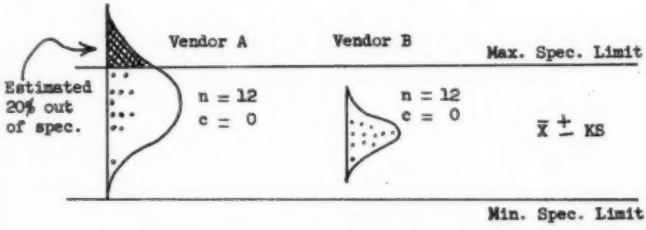
For example, environmental data was received from two different vendors on a humidity test. The results for both on an attribute basis were the same. The sample size tested 12, and zero defects noted. Under these results either vendor was capable of manufacturing a product within the humidity test requirements. A review of the actual readings of the humidity data by the Quality Analysis group produced an entirely different conclusion.

HOW MANY DOLLARS ARE THERE IN INCOMING INSPECTION COSTS?



Notice that neither vendor has any readings outside the specification requirements. By some simple calculation for $\bar{X}$ and sigma, several features distinguish one vendor from the other as to manufacturing capabilities. For example, Vendor B has a better average value and uniformity of manufacture and a forecast of the lot quality shows that the individual values of the lot are also within the specification requirements as calculated by $\bar{X} \pm KS$. where K is specified as 95% probability that 99% of the distribution will fall between the specification limits.

In the analysis of Vendor A, the average value is above nominal and the sigma is so large that 20% of the units in the lot from which this sample was taken would be outside of the upper specification limit.



Results of course are obvious. Vendor A was notified of the results and in turn Vendor B was given preference in future orders. Cost savings on such simple statistical techniques are immeasureable.

The next step was to develop some method of plotting environmental information using a common denominator to permit similar as well as indicated parameters to be plotted on the same chart.

For example, the MIL-R-11 resistor is designated as a fixed carbon resistor. MIL-R-11 resistors are available in many resistance values and tolerances. It would take a long period of time to obtain a history of environmental test results on any one resistance value. Therefore, it would be to our advantage to combine all environmental information on MIL-R-11 fixed carbon resistors. On the same chart it is possible to plot the test results of resistors of 10,000 ohms $\pm$ 5% with 5,000 ohms $\pm$ 1% with 500,000 ohms $\pm$ 10%, etc. The end result will be to evaluate the environmental test data of the MIL-R-11 resistor. The denominator chosen was the number of standard deviations the average value of the parameter is away from the upper or lower specification limit.

HOW MANY DOLLARS ARE THERE IN INCOMING INSPECTION COSTS?

This unit (X/σ), readily permits determination of the percent rejects to be expected even though all samples examined are within the specification limits.

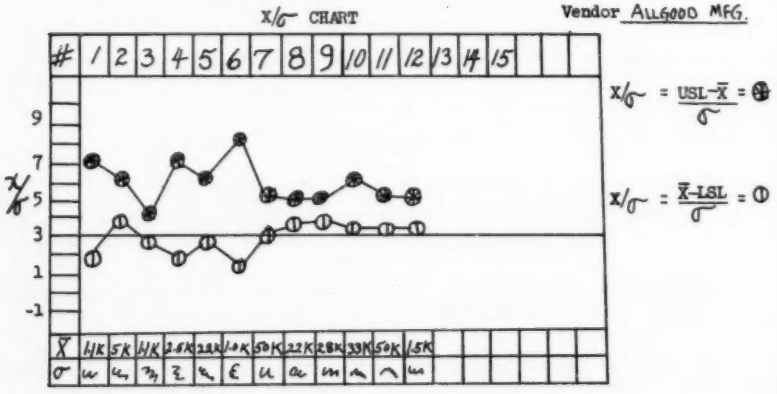
X/σ is equal to the number of standard deviations which the sample average is away from the upper or lower specification limits. The sigma is calculated from the individual values of the sample.

The $\frac{X}{\sigma} = 3$ line permits a ready reference point. All values of $\frac{X}{\sigma}$ of 3 or greater are indicative of good product; the larger the value the better, provided the sample average is above the nominal for the upper specification limit and below the nominal for the lower specification limit. Values less than 3, or approaching a negative value, are indicative of more and more discrepant material.

Values of $\frac{X}{\sigma}$ greater than 10 are not shown, but they are plotted as 10, since 10 standard deviations away from the specification limit will indicate that the sample sigma is very small.

The plots normally will be for a single vendor on a multiple chart form.

Environmental Test Data - MIL-R-11 Resistors
Resistance Check after Humidity Test



By interpreting similar charts, here is what Quality Control can say to the Purchasing Department.

1) When the symbol ⊕ is consistently above the symbol ① line, the product is consistently below the nominal of the specification requirement.

2) If the two symbols tend to cross one another, then the vendor is close to the nominal value of the specification.

HOW MANY DOLLARS ARE THERE IN INCOMING INSPECTION COSTS?

3) The wider the separation between symbols, the farther the vendor is away from nominal, i.e., in the above chart, lots 1, 4, 5, and 6 show the vendor too far below the nominal value.

4) In the above chart, lots 3, 8 and 9 are nearest to the nominal value of the specification.

5) The larger the sum of the two symbols, the more uniform variability pattern of the vendor.

6) This chart can distinguish between process capabilities problems and/or problems due to average levels of product, i.e., if the vendor has a point below the 3 line and the sum of both symbols is greater than 6, the problem is one of average level and not process capabilities. The above chart shows the vendor having a problem with average level of product.

7) Determine how good a shipment is with reference to the component specification.

8) Determine whether there have been any changes between shipments in the quality level of a vendor or part number.

9) Compare vendors so that a vendor rating may be developed.

There are several ways the environmental data may be used for inclusion in a vendor rating. Perhaps for a start, the vendor could be rated as acceptable or not acceptable based on the environmental data as a supplement to the process average rating plan.

Purchasing must have immediate and dependable quality and reliability data upon which to base a procurement decision.

TOTAL QUALITY CONTROL THROUGH RELIABILITY

Frank M. Gryna, Jr.
TM-76B Program Reliability Administrator
The Martin Company, Baltimore, Maryland

In the past, many quality control programs have concentrated on the manufacturing phase of operations. Recently, however, the increased emphasis on reliability has made it necessary to expand the traditional quality control approach to cover design, procurement, manufacturing, test, and field operations to assure operational reliability -- i.e. continued proper performance in the field. The need for a comprehensive program is illustrated by the results of an analysis of 815 equipment failures (1):

43% of the failures were due to design
20% of the failures were due to manufacturing
30% of the failures were due to field conditions

(The remaining 7% were due to parts wearing out and other unknown causes.)

This paper describes how Martin-Baltimore has expanded its quality control concept into a comprehensive reliability program. Included is a discussion of reliability policies, reliability organization, and a method of defining a reliability program.

RELIABILITY POLICIES

Our program is based on these policies:

1. Reliability is an element of product performance distinct from capability. Capability is the ability of a product to perform a function. Reliability is the ability to perform the function repeatedly or when called upon to do so.

2. A reliability program must start in the proposal phase of a project and continue throughout design and development, production, test, field evaluation, and service use. This means that the program cannot be restricted to any one organizational unit but must cover all units that effect the final field reliability.

3. Adequate funds must be provided for a reliability program and these must be determined during the proposal phase. This means that a complete reliability program must be developed in sufficient detail during the proposal effort to permit adequate costing.

4. The execution of a reliability program involves both technical tasks and a management task. The technical tasks consist of the efforts to design reliability into the product and maintain this reliability throughout production and field use of the product with minimum degradation. The management task consists of integrating all of the technical efforts, and controlling these efforts to assure that all necessary steps are being taken to achieve the required reliability.

5. Reliability results can only be achieved by actions taken by the line organization -- the designer, the production man, the procurement man, etc. The reliability specialist provides guidance and assistance to the line personnel in executing their fundamental reliability tasks.

6. The program for each project must provide a written plan and must specify responsibility, procedures, and schedules.

7. The program must include controls which will detect and report to management all deviations between plans and actual performance.

8. The program must include suppliers as well as internal Martin operations.

9. The overall integration and evaluation of the reliability program must be performed by an organization which is independent of those who are responsible for taking the detailed steps necessary to achieve the required reliability.

These policies are the foundation of the working reliability program described in the following paragraphs.

ORGANIZATION FOR RELIABILITY

The products of Martin-Baltimore range from electronic systems such as a flight control system to complete weapons systems such as the TM-76B Mace missile. Martin-Baltimore is organized both on a functional division basis and a program organization basis; i.e. the various functional divisions supply people to individual projects.

The Vice-President and General Manager of Martin-Baltimore has the overall responsibility for the capability and reliability of the product. He has delegated to the Quality Division the responsibility for establishing, maintaining, and assuring the adequacy of an integrated reliability effort within Martin-Baltimore, and on its individual programs.

On each program, the Program Manager is responsible for capability and reliability. As the reliability program is both preventative and corrective in nature, we have found it useful to setup the project reliability organization with this in mind. Fundamental to the organizational approach is the definition of responsibilities for each of the functional divisions with the overall responsibility for the entire program resting with the Program Reliability Administrator. A Reliability Control Team (Figure 1) is established to enable the administrator to effectively develop and coordinate the individual divisional reliability tasks. This team consists of reliability specialists from each division on the program. Each member of the team is responsible for planning, directing, and controlling the reliability tasks defined in the program for his division. This means that he must define the tasks in writing in the reliability program document, execute and/or provide technical assistance to those performing the tasks, and continuously monitor the reliability tasks for his division. The Reliability Control Team devotes most of its efforts to the preventive aspects of the program. Each project also has a Corrective Action Team (Figure 1) which is responsible for defining problem areas based on actual equipment failures and following up until corrective action is put into effect and its adequacy verified. We have found that the fixing of responsibility on a Program Reliability Administrator and the use of both a Reliability Control Team and a Corrective Action Team provide continuous emphasis on both the preventive and corrective phases of the program. Within the Quality Division a separate department -- Reliability and Quality Assurance -- exercises the divisions function of assuring the adequacy of the overall company reliability effort, and provides the Program Reliability Administrator to act as Chairman of the Reliability Control Team.

Since a large portion of reliability activity is accomplished during the design stage, the Engineering Division has established a special reliability organization to assure implementation of the engineering portion of the program.

MRM - THE MARTIN RELIABILITY MANAGEMENT PROGRAM

We have found the following steps to be important in developing a reliability program for a specific project:

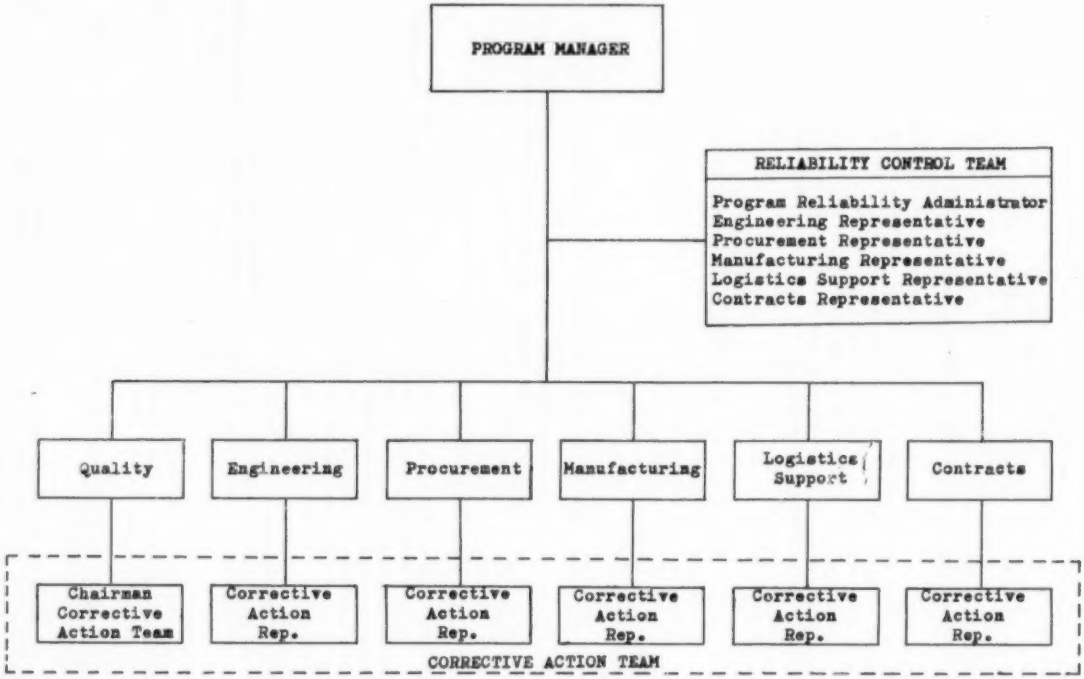

PROGRAM MANAGER
RELIABILITY CONTROL TEAM
Program Reliability Administrator
Engineering Representative
Procurement Representative
Manufacturing Representative
Logistics Support Representative
Contracts Representative
Quality
Engineering
Procurement
Manufacturing
Logistics Support
Contracts
Chairman Corrective Action Team
Corrective Action Rep.
Corrective Action Rep.
Corrective Action Rep.
Corrective Action Rep.
Corrective Action Rep.
CORRECTIVE ACTION TEAM

FIGURE I

1. Define the reliability objective in specific terms:

From the viewpoint of defining objectives, to say that we want "high reliability" is as meaningless as saying we want "low cost". Industrial engineers insist upon cost budgets to define cost objectives. In reliability we insist upon a numerical definition of the reliability objective, e.g., a transmitter must have a 99% reliability for a two hour operating period. This is not always a simple matter and for weapons systems may involve military operations research techniques to determine the reliability objective for a specific equipment.

2. Define each of the steps to be taken in the program:

Generalities such as "use the best components available" or "maintain high quality workmanship" are meaningless in defining a reliability program. The steps must be defined in specific terms in order to permit adequate costing and develop tight controls for each step. I have found that a good criterion for a specific statement of a step in the program is whether or not the step can be costed. Costing is important when there are steps in the reliability program for a specific project that require efforts above those in the general company reliability program.

3. Establish the responsibility for each step in the program:

The responsibility for each step must be established and assigned to a single unit of the organization. This may require that a step be divided into several elements in order to assign responsibility to a single unit.

4. Define the other organizational units concerned with each step:

Pinpointing the responsibility for each step is basic, but most steps involve several segments of the organization. Defining the units concerned with each step shows the coordination required.

5. Provide adequate funds for each step in the program:

Many steps in a reliability program are a part of the normal company operating procedure and do not require additional funds. However, on a specific project, certain efforts may require additional funds. (An extensive reliability testing program is a typical example.) Sometimes, however, all that is required is a change in the distribution of funds from a later phase of the program to an earlier phase. Defining the steps in specific terms makes it comparatively easy to estimate the funds required. However, all of this must be done in the proposal phase of a contract to assure that the submitted price contains funds for reliability either separately identified or included as part of the divisional efforts -- depending upon customer preference. You can well imagine the difficulty in establishing a complete reliability program after a contract is received if adequate funds have not been included in the submitted price. Finally, it should be emphasized that funds provided for reliability are an investment that will pay off in terms of lower total program costs.

6. Prepare written procedures detailing how each step in the program will be performed:

As one step in assuring implementation, written procedures must be prepared and distributed in advance to all those who have a role in the specific program step.

7. Establish time schedules for the preparation of written procedures and completion of the program step:

Reliability is part of the overall project effort and must conform to project time schedules. To assure this, time schedules must be set for the preparation of written procedures and the completion of the actual tasks.

8. The "control" aspect of reliability:

Martin makes a distinction between reliability planning and reliability control. The most brilliant plans are academic unless specific measures are taken to control the execution of the plans. In the Martin program, the listing of tasks is regarded as a set of plans. We feel that this listing is a unique way of definitizing the program. However, for each task, we also state the specific controls that will be used to make sure the job is done and done properly. These controls are shown in Table 1. Reliability controls, then, are not just a few isolated tasks -- they are specific measures taken to assure that each and every step in the program is executed.

The steps listed above result in a document we call the MRM -- The Martin Reliability Management Program. The MRM program for a project usually contains anywhere from 30 to 60 steps. Table 1 shows the actual format of an MRM and lists a few representative steps from an actual program. The MRM is the Martin approach for changing reliability from a vague concept to a definitive program.

An MRM program is developed for each project - and issued as an official project directive. We feel that the MRM document is useful not only to Martin but also to our customers. For Martin, it is not only a means of control but during the proposal phase it also provides a basis for costing the reliability effort to assure that sufficient funds have been allotted to execute the program. For our customers, it provides a description of what specific steps we will take in a reliability program and how these steps will be controlled. We welcome requests for bid that require a detailed description of a reliability program.

Martin experience with military products has emphasized the need for reliability programs that start in the initial design phase and extend throughout all program phases including field use of the product. The traditional emphasis of quality control during manufacturing has been expanded to provide an attack on all of the causes of low reliability. Reliability problems force a broadening of traditional quality control to the concept of total quality control.

1. Greenidge, R. M. C., "The Case of Reliability Versus Defective Components Et Al", Electronic Applications Reliability Review, Issue No. 1, 1953, p. 12.

TABLE I - Excerpts from the Martin Reliability Management (MRM) Program for the_____Project

Task	Primary Responsibility	Other Functions Concerned	Procedure Approved		Controls	Schedule	
			Date	By		Start	Compl.
Perform reliability predictions based circuit and component analyses and use of reliability handbook. Compare to reliability budgets. If any design revisions are made, review reliability prediction and revise as required	Design Engr.	Quality	Completed	Tech. Dir.	1. Reliability predictions to be included on page and line schedule. 2. Analyses to be reviewed by group engineer. 3. All analyses must be submitted to reliability engineer for his approval.	7/1/58	10/1/59
Prepare major subcontract specs. with adequate definition of numerical reliability requirements and objectives, reliability program, tests, reports and related requirements, such as performance, environment and operational, maintenance and checkout and support concepts.	Contracts (Specifications)	Engrg., Quality, Procurement	9/15/58	Contracts Mgr.	1. Specifications must be reviewed by Reliability Control Team and approved by Contracts Manager before issuance to vendor.	10/1/58	2/1/59
Include reliability requirements in evaluation and selection of vendors and subcontractors.	Procurement	Engrg., Quality	8/15/58	Procurement Mgr.	1. Vendor selection to be subject to review by Reliability Control Team.	9/1/58	10/1/59
Review all designs to determine component part and assembly limitations with respect to reliability requirements using economical processes and procedures.	Elect. Mfg.	Quality	8/15/58	Mfg. Mgr.	1. All drawings must be reviewed by Manufacturing representative prior to engineering release.	9/1/58	1/1/59

TABLE I (Cont'd)

Task	Primary Responsibility	Other Functions Concerned	Procedure Approved		Controls	Schedule	
			Date	By		Start	Compl.
Establish failure reporting and problem resolution system from first development test through service use. a. Establish and maintain failure reporting in test areas and service use. b. Establish and maintain failure reporting in selected vendors plants. . . .	Quality	Engrg., Mfg., Proc., Logistics Support	2/15/59	Quality Mgr.	1. Assign specific Reliability and Assurance engineer and Corrective Action Team to find problems and assure correction. 2. Monitor reporting through assigned personnel in reporting areas.	3/1/59	Compl. of Service Use
Verify handbooks (including use by anticipated skill levels) on development units. Include both plant and field verification. a. Revise handbooks and/or redesign as needed. . . .	Logistics Support	Engrg., Quality, Mfg.	10/15/59	Logistics Support Mgr.	1. Establish index of handbook errors per unit of preparation time. 2. Corrections must be approved by Logistics Support Manager.	11/1/59	2/1/60

PROGRESS IN PROFESSIONALISM

W. H. Lewis
General Electric Co., New York, N. Y.

The growth of the American Society for Quality Control in only 14 years to a membership of 12,000 attests to the constantly increasing interest in the field of quality control. Pick up any metropolitan newspaper today and look in the classified advertising section. Companies across the nation are seeking competent engineers who can specialize in controlling and improving product quality. This fact serves to emphasize how highly technical the work has become in recent years. Indeed, it presents a challenge to the American Society for Quality Control in fulfilling the needs of American Industry with professional quality control engineers. More than ever before, businesses recognize product quality as an essential factor in meeting increasing competition and maintaining a profitable position. In fact, our very survival as a nation is linked with the ability of industry to produce products that will reliably fulfill their function. I need not dwell here on the necessity for good quality and reliability in our missiles and other military equipment.

The American Society for Quality Control has accepted the challenge to help provide industry with such professionally qualified engineers. How well it is dispatched will determine its future success and the success of its members in the quality control engineering profession. As a Society,how far have we come? As many of you know, the Professional Development Program has been active for approximately two years. At the 1959 Convention of the American Society for Quality Control, a series of papers "Professionalism in Quality Control" attracted much attention. Several of the papers in this series have subsequently appeared in Industrial Quality Control. One of these papers in particular served to summarize the professional development program up to that time. This was presented by Mr. Leon Bass under the title "What ASQC is Doing for the Quality Control Engineer." It appeared in the December 1959 issue of Industrial Quality Control. Also of interest in this regard was the final report of the Special Committee on Professional Development which appeared in the October 1959 issue of Industrial Quality Control. These reports covered the activities with respect to delineation of the work of the professional quality control engineer, state licensure, development of college curricula, management relationships and professional development at section level. The purpose of this paper is to inform you of the progress that has been made in these fields of activity during the past year.

One of these basic programs, and a very important one, was undertaken by the Work Elements Committee. The Chairman of that Committee, Mr. Robert Schin, is presenting the work of that committee in a paper at this Convention as it has progressed to date. Membership of that committee was drawn from various industries, universities and government positions. The nine members of the committee have outlined the kinds of work a quality control engineer should perform. And, after some refinements, these elements of work for professional quality control engineering will be published in Industrial Quality Control for review and comment by our membership. When these work elements are grouped into specific areas of activity, the various technologies of Quality Control begin to emerge. These then begin to define the larger areas of specialized knowledge that are required to do a professional quality control job. Dr. Feigenbaum is presenting a paper at this Convention which further develops the specialized technologies in Quality Control. As will presently be shown, clear delineation of these technologies is essential to the development of quality control to a professional level.

In another area, a committee on Professional Licensing was established under the Chairmanship of Mr. John G. Rutherford. Work here has been largely exploratory. State boards for licensing engineers are limiting the number of classifications under which examinations are given to those where specialized

fields of knowledge are required. This fact serves to underline again the need to define clearly the technologies of quality control. Furthermore, established curricula in colleges and universities should be in place to provide opportunity for acquiring such specialized fields of technical knowledge.

The Committee on Professional Licensing has succeeded in getting the United States Civil Service Commission to classify Quality Control in its classification standards. Some discussions are going on concerning the major classification under which quality control engineering should appear. It is tentatively listed under industrial engineering classification. Other probabilities being considered are: the Commodity Quality Control, Inspection and Grading Group; or the General Engineering Series. The latter classification would seem to place quality control engineering more in character - particularly with increased emphasis on the technical knowledge required to generate and assure product quality.

In view of the problems encountered in getting quality control recognized as a specialized field of engineering - particularly with respect to State licensure - it becomes increasingly evident that we must clearly delineate the areas of specialized knowledge that are involved. As was pointed out earlier, grouping of the work elements helps to define such areas. We must go even further and define the disciplines and courses of study required to support such technology. For example, if we consider process control as a technology, the supporting disciplines would include industrial statistics with such studies as designed experiments, capability studies, and control charts. Another discipline might be electronics with subjects in measurement, instrumentation and closed loop feedback for process control. It now becomes quite evident that we must work in the direction of developing suitable curricula to serve as a basis for the education of quality control engineers. To this end, the Committee on College Relations has been activated for the purpose of working with colleges and universities in developing and installing such curricula.

What are our chances for success in this area? I think we must first understand the position now being taken by the majority of colleges and universities. As you may surmise, these institutions are constantly being approached by various groups to install curricula covering each group's specialized interests. To yield to even a small percentage of such requests is beyond the limits of faculty, facilities and budgets. Colleges and universities are taking the position that only general courses of study can be considered for the undergraduate schools. This means that Engineering Schools are limiting their undergraduate curricula to Aeronautical, Chemical, Civil, Electrical, Industrial, and Mechanical Engineering. Any specialization beyond this is set over into the graduate school. This may be an advantage - particularly in view of some recent trends in advanced degree work. I was surprised to learn the other day that over half of the graduate students of a large Eastern university are employed in industry and are attending school at night. We already see reputable universities and colleges installing some courses that would be a part of a quality control curriculum. It is only a matter of time until suitable curricula will be available, but our College Relations Committee can be instrumental in hastening the development along the most productive lines.

Some basic courses are also becoming available in the undergraduate school. These may interest undergraduate engineers in specializing in quality control during graduate study. It may be only a matter of a few years until several universities will have a quality control curriculum; that will provide a young man with an opportunity to train thoroughly for the quality control engineering profession.

What opportunities can be provided for the established man who is in industry and seeking to grow professionally? Some of these men may be situated near universities and could avail themselves of courses offered in the graduate school or university extension divisions as they become available. Another interesting possibility that is being given considerable thought and study is the founding of a Quality Control Institute under the auspices of ASQC. This might be patterned after the American Banking Institute which has enjoyed sixty years

of successful work in the field of banking education and training. Such an institute provides several complete curricula and issues a certificate in recognition of the satisfactory completion of each. This goes far beyond the sporadic two-day conference or eight-day course that has been characteristic of quality control educational efforts in the past. While these short courses have filled a need, they can hardly be considered as having created professional quality control engineers.

I have dwelled on the matter of education at considerable length because of its importance. It offers an opportunity for the individual to gain knowledge necessary to become a true professional quality control engineer. Of course, education by itself is not enough. Its value must manifest itself in the performance of the individual as he serves in his profession. Exactly the same thing can be said of experience. It must manifest itself in increased skill and performance to be considered valuable experience.

Let us say that satisfactory completion of a certain educational curriculum and ten years of experience in a certain capacity does equip the average individual to meet a given standard of performance. Then, industrial managers recognizing this, can more easily give greater recognition to the individual desiring to advance in his chosen profession. That is the reason the Professional Development Council includes a Management Relations Committee. Through this committee, we are seeking to reach our "market." We want the customer to know the "qualities" of the product available to him. In other words, we are seeking recognition of the professional standards for quality control when and as they are established. In this regard, it is extremely important that we preserve our integrity. False representation would serve only to wreck quickly and completely the whole professional development program. When we have professional people qualifying under established standards, who can deliver a high level of performance, we should then speak of them. No doubt there are some who now qualify as professional quality control engineers, but the number is not sufficiently great to gain universal recognition among industrial managers - especially when we consider there are over 46,000 industrial establishments producing 60% of the Nation's products. We can, however, be establishing lines of communication and be recognized for the present state of the art. Mr. L. A. Seder and his Committee on Management Relations will be concentrating on this activity during the coming months.

In summary then, the work of the various ASQC committees engaged on the Professional Development Program have shown the course at national level to be clear:

1. The technological fields of quality control must be clearly defined.
2. There must be curricula established that will provide training in these technologies.
3. Professional standards must be established and individuals must be examined and certified as having met a given standard.
4. These standards must be recognized by industrial managers for their value in terms of professional performance.
5. A sufficient number of individuals must attain these professional standards and have an impact on the industrial community to gain recognition for quality control engineering as a profession.

In connection with this last point, the appeal is directly to the individual. In the final analysis quality control engineering will be recognized as a profession because a large number of you as individuals earn the right to be called professional quality control engineers.

What then, as individuals, can you do to earn professional recognition?

1. You can identify the bodies of knowledge necessary for quality problem solving in the particular industry in which you are engaged.

2. You can obtain this knowledge from several sources: university

extension courses, ASQC sponsored courses, self-improvement programs such as professional reading and study.

3. You can, through application of this knowledge, become skillful in solving the product quality problems that present themselves.

Where several of you with like ambitions can associate with each other in the local ASQC section, mutual support can be realized. The Section Liaison Committee under Mr.Gordon Thompson is encouraging local sections to form professional development committees. These local groups can gain support for requests to local colleges and universities in providing necessary courses of study. These many local efforts can materially help achieve establishment of required curricula on a national scale. It is by such processes that needed curricula will become established. And, it is through such a program of self-development on the part of each of you as individuals that the American Society for Quality Control will be able to accept the challenge in providing American Industry with competent people - people who can make outstanding contributions to the public, to business growth and profitability, and to the Nation.

THE ANALYSIS OF VARIATIONS FOR CERTAIN PHYSICAL PROPERTIES OF COMBED COTTON YARNS

D. S. Hamby, W. C. Stuckey, B. Gast, School of Textiles
R. J. Hader, Institute of Statistics
North Carolina State College
Raleigh, North Carolina

INTRODUCTION

For a number of years the Faculty and Staff of the School of Textiles at N. C. State College have been asked by the industry to test different textile materials and to comment on the resulting quality levels. A good portion of these requests involved yarn, and a search of the literature did not reveal any suitable references of such material. The information was not available because some particular references pertained only to a specific topic or because the advance in testing techniques have developed at such a rate that the literature is not complete on some subjects.

Since this type of information was not available in suitable form, it was decided to conduct a study designed to give information as to the levels of the quality measurements normally used in evaluating yarns.

Although every effort was made to collect samples that were representative of the combed yarn industry, no claims are made by the authors that the results represent standards as such for the variables that have been measured.

MATERIALS STUDIED

A total of 61 combed yarns, ranging in count from 16/1 to 120/1 were sampled from 32 different mills. For each yarn number 25 spinning bobbins were selected at random. In this first study no effort was made to separate between-side, between-frame and between-shift variation.

The mills that participated were selected according to their reputation in the trade. Some mills considered to be producing the best yarn in the industry were selected and others considered to be producing some of the lowest quality were also selected. Most of the mills, however, were in the so-called average quality class.

EXPERIMENTAL PROCEDURE

Table I summarizes by type of test the number of bobbins, k, and the number of measurements per bobbin, n, for each yarn.

For skein strength measurements, 120-yard skeins were broken on a 150-300 pound capacity Scott Model J tester. All single-end strength measurements were made on the Uster Dynamometer with a speed setting of 5.

The yarn number was measured on an Alfred Suter Direct Numbering Balance using the skeins from strength determinations.

TABLE I
AVERAGE VALUES FOR RESULTS OF ANALYSIS OF VARIATION OF COMBED COTTON YARNS

Yarn Property	k	n	Average*					
			CV_o	CV_b	CV_w	CV_r	CV_{fd}	Dev.
Skein Strength Pounds	25	4	6.0	5.0	3.8	3.6		
Single End Strength Grams	10	20					10.8	
Yarn Number English System	25	4	2.7	2.2	1.6	1.5		2.4
Twist Turns Per Inch	25	4	5.0	3.1	4.4	4.4		
Uniformity Per Cent U	10	1						

Legend:
k = Number of bobbins tested for each of the 61 combed yarns
n = Number of tests per bobbin
CV_o = Per cent coefficient of over-all variation
CV_b = Per cent coefficient of variation between packages
CV_w = Per cent coefficient of variation within packages
CV_r = Per cent coefficient of variation (Range Method)
CV_{fd} = Per cent coefficient of variation (Frequency Distribution Method)
Dev. = Per cent deviation from specified number

* Each value for skein strength, yarn number and twist represents 6100 measurements. Each value for single end strength represents 12,200 individual measurements.

TABLE II
SOURCES OF VARIATIONS IN YARN NUMBER AND SKEIN STRENGTH FOR COMBED COTTON YARNS

Source	Per Cent of Total Variation	
	Yarn Number	Skein Strength
Within Bobbin	33	34
Between Bobbin	44	47
Between Sides of Spinning Frames	10	5
Between Frames	13	14

The Uster Mean Linear Integrator was used in conjunction with a Uster Model B Evenness Tester to measure the per cent linear deviation.

The 10-inch untwist-twist method was used to determine the twist per inch using the twist tester manufactured by the U. S. Testing Company.

DATA ANALYSIS

The main portion of the data analysis consisted of application of the analysis of variance technique to study the "between bobbin" and "within bobbin" variation for each of the three properties: yarn number, skein strength and twist. A separate analysis of variance was performed on each of these properties for each of the 61 different yarns. In each case there were twenty-five bobbins and four tests per bobbin. The general form of the analysis of variance is:

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	Expected Mean Square
Between Bobbins	24	$4\sum_i (\bar{x}_i - \bar{\bar{x}})^2$	M_b	$\sigma_w^2 + 4\sigma_b^2$
Within Bobbins	75	$\sum_i \sum_j (x_{ij} - \bar{x}_i)^2$	M_w	σ_w^2
Total	99	$\sum\sum (x_{ij} - \bar{\bar{x}})^2$		

In this analysis the total variation around the grand mean, $\sum\sum (x_{ij} - \bar{\bar{x}})^2$, is broken up into two parts, the first based on differences of bobbin means from the grand mean, $4\sum (\bar{x}_i - \bar{\bar{x}})^2$ and the second a pooling of the within bobbin means, $\sum_i \sum_j (x_{ij} - \bar{x}_i)^2$. These sums of squares are then divided by their respective degrees of freedom to yield the mean squares M_b and M_w. By statistical theory it may be shown that the expected or average value of these two mean squares are $\sigma_w^2 + 4\sigma_b^2$ and σ_w^2 respectively. The symbol σ_w represents the standard deviation of the "within bobbin" component of variation and σ_b represents the standard deviation of the "between bobbin" component. Once M_b and M_w have been calculated numerically the values of σ_b^2 and σ_w^2 are readily found by solving the equations

$$M_b = \sigma_w^2 + 4\sigma_b^2$$

$$M_w = \sigma_w^2$$

For purposes of presentation of results all values of σ_w and σ_b were converted to coefficients of variation by dividing by the mean, $\bar{\bar{x}}$, and multiplying by 100. It was also of interest to combine the two components of variation into an over-all coefficient of variation which is found from the relation

$$CV_o^2 = CV_w^2 + CV_b^2$$

These are illustrated by an example shown under Sample Calculations.

The "within bobbin" variance may more conveniently be estimated by the range method using $\sigma = \bar{R}/d_2$ where $\bar{R}$ is the average "within bobbin" difference between largest and smallest values and d_2 is conversion constant dependent on the number of tests within each bobbin.* For four tests $d_2 = 2.059$.

The various coefficients of variation were subjected to further study consisting of frequency distributions for each type and of attempts to correlate the various types with each other and also with additional variables such as, for example, the CV_o for skein strength vs. the coefficient of variation for single-end tests. This will be discussed in more detail in the next section.

RESULTS AND DISCUSSION

The difference in level for the coefficients of variations for over-all, between package and within package for skein strength are obvious in Figure 1. The over-all coefficient of variation (CV_o) averaged 6.0 per cent and included within the within-bobbin, between spindles, between sides and between frames variation. The between package variation (CV_b) averaged 5.0 per cent which, as would be expected, was slightly less than the over-all. The within-package variation (CV_w) averaged 3.8 per cent. Since the range method of calculating the standard deviation is commonly used in quality control work, this method was also used in addition to the standard analysis of variance technique. As seen in Figure 1, as well as other figures, the coefficient of variation based on the range method (CV_r) is in close agreement with the CV_w. Although the calculations were not made in this study, the over-all coefficient of variation may also be estimated from ranges. This method is explained by Enrick (1) and has been in use for a number of years.

Figure 2 shows the distribution and level for the coefficient of variation (CV_{fd}) of the single-end strength. These values were calculated from the frequency distribution as accumulated on the Uster Dynamometer. Each entry represents the distribution of 20 measurements each for 10 bobbins.

The variation in yarn number is shown in Figure 3. The CV_o averaged 2.7 and with the exception of one yarn ranged from 0.5 per cent to 4.5 per cent. The CV_b averaged 2.2 and the CV_w averaged 1.6. As was the case with the skein strength, the CV_r and CV_w were in close agreement. Each measurement entered in this figure represents four measurements each from 25 bobbins.

A comparison is shown in Figure 4 between over-all coefficients of variation for the combed cotton yarns used in this study and for worsted yarns.

The details of the worsted yarn study were presented by Bornet (2) and a comparison between the two types of yarns show a remarkable degree of similarity in this particular property.

Since the original study of yarn number variation was not designed to permit a complete analysis of the variation in yarn number and strength, subsequent studies (references 3 and 4) were made to determine the magnitude of the influence of spindles, sides and frames. A summary of these studies is shown in Table II. It is interesting to note the close agreement between yarn number and strength when comparing the sources of variation. Beckwith (5) reported that 32 per cent of the variation was within cones and 68 per cent between cones. The results of this study confirm these findings.

In view of the inter-relationship of certain physical properties of yarns the basic data from this study was further analyzed in an effort to determine the degree of some of these relationships. Figure 6 shows the correlation between skein strength and single-end strength. This, as well

---

* ASTM Manual on Quality Control of Materials, Table B2.

as subsequent data, shows excellent correlation ($r = 0.996$) between these two properties. The equation $Y = -6.6917 + 0.3011X$ is for X in units of grams. For X in pounds the equation becomes $Y = -6.917 + 136.58X$.

The relationship between the Mean Linear Deviation, Per Cent U, as measured by the Uster Linear Integrator, and the logarithm of the yarn number proved to be most interesting. As shown in Figure 7 there is a distinct difference between the combed and carded yarn. Only one combed yarn exceeded the limits on the high side and only two carded yarns entered the combed yarn area. Subsequent combed yarn data not shown on the chart all fell between the two limits. A comparison between these limits and the Uster Standards will show the limits in this figure to be somewhat lower than the Uster Standards.

Figure 8 is a cumulative frequency distribution showing that slightly less than 80 per cent of the mills tested were within the American Society for Testing Materials Tolerance of 3 per cent. Figure 9 shows the distribution above and below target number.

There appears to be no relationship between yarn number and per cent deviation from specified count or yarn number and CV_o for yarn number. Coefficients of variation were not calculated, however, by examining Figures 10 and 11 no trends are apparent.

There is some correlation ($r = 0.83$) between the coefficient of variation for single-end strength (CV_{fd}) and the Mean Linear Deviation, Per Cent U. This plot is shown in Figure 12.

No statistical analysis was made of the data in Figure 13, however, by visual examination it is obvious that there is no trend or pattern between CV_o of skein strength and CV_{fd} for single-end strength.

Figure 14 shows that there is fair to poor correlation for over-all coefficients of variation for yarn number and skein strength and Figure 15 shows fair to poor correlation for CV_w for yarn number and CV_w for skein strength.

The CV_{fd} for single-end strength and the CV_o for turns per inch exhibited very poor correlation ($r = 0.38$).

PRACTICAL APPLICATIONS OF RESULTS

In order to illustrate the practical application of the results and conclusions of this study, suppose that the quality control program for a mill develops the following information: for skein strength a CV_o of 6.54 per cent and a CV_w of 3.37 per cent; for yarn number a CV_o of 2.70 per cent, a CV_w of 1.29 per cent and a deviation from specified count of 2.15 per cent. By examining Table I, the value 6.54 is found to be higher than the average of 6.0; also by inspecting the bottom diagram in Figure 1, the value 6.54 is in the upper portion of the diagram. This means that if this yarn is to be considered as an average to good quality yarn, this variation should be reduced. The CV_w of 3.39, is better than average. Since the variation between bobbins is high or worse than average, effort should be directed toward reducing the variation in strength from bobbin to bobbin.

The per cent deviation of the yarn number from the nominal or specified count of 2.15 per cent is below the average of 2.4 as shown in Table I and also is below the 3 per cent specified by the American Society for Testing Materials. The CV_o of 2.70 is the same as the average shown in Table I. The CV_w of 1.29 is below the average of 1.6 and the CV_b of 2.46 is higher than the average value of 2.2.

It should be remembered when using this type of analysis and interpretation that the coefficient of variation will vary to some extent as the result of sampling variation. If there is a doubtful case as to the significance of the level of a variable further statistical tests should be made on the data before any definite decision is made.

ACKNOWLEDGEMENTS

The authors are indebted to Dr. A. H. Grandage of the Department of Experimental Statistics at North Carolina State College for his cooperation and advice in the analysis of the test data. Professor E. B. Grover, Head of the Department of Fiber and Yarn Technology, School of Textiles, N. C. State College, was most helpful during the course of this study with numerous suggestions and advice. Hassan Makhlouf, Ahmed Baz, Mahmoud Ibrahim, and Nabil Costandy, graduate students of the School of Textiles, contributed to the study by performing special calculations during the preparation of the final report.

LIST OF REFERENCES

1. Enrick, N. L., Quality Control Through Statistical Methods, Modern Publishing Corporation, 1954.

2. Bornet, George M., "Canadian Yarn Survey - New Approach to Causes of Count Variation," Textile Quality Control Papers, Vol. 5, pp. 39-72, 1958.

3. Baz, Ahmed, "Design of Sampling Plans for Skein Strength of Yarn Based on the Analysis of Variations." Thesis for Master of Science Degree, School of Textiles, N. C. State College, Raleigh, North Carolina, 1959.

4. Ibrahim, Mahmoud, "Design of Sampling Plans for Yarn Number Measurements Based on the Analysis of Variations." Thesis for Master of Science Degree, School of Textiles, N. C. State College, Raleigh, North Carolina. 1959.

5. Beckwith, O. P., "How Many Cones Shall We Test," Textile Industries, Vol. 122, No. 2, pp. 104-106, 1958.

6. Grover, E. B. and D. S. Hamby, Handbook for Textile Testing and Quality Control, N. C. State College, 1958.

7. Hader, R. J., Introduction to Statistics for Engineers, N. C. State College, 1957.

SAMPLE CALCULATIONS

The conclusions and body of this report have been written in such a manner as to facilitate easy interpretation and use by management and other supervisory personnel. For the readers that are interested in the statistical techniques used in the analysis of the data, the following sample calculations are presented.

Total variance* $= \Sigma(X - \bar{\bar{X}})^2 = 20600.6$

Variance within bobbins* $= \Sigma(X_i - \bar{X}_i)^2 = 6750.0$

Variance between bobbins* $= N\Sigma(\bar{X}_i - \bar{\bar{X}})^2 = 4(3462.6) = 13,850$

	Sum of Squares	Degrees of Freedom	Mean Squares
Within bobbins	6750.0	75	90.0 $= \sigma^2_w$
Between bobbins	13850.0	24	577.1 $= \sigma^2_w + n\sigma^2_b$
Total	20600.0	99	

$\sigma^2_w + 4\sigma^2_b = 577.1$

$\sigma^2_w = 90.0$

$\sigma^2_b = \frac{577.1 - 90}{4} = 121.78$

$CV_w = \sqrt{\frac{90}{248.4}} \times 100 = 3.81\%$

$CV_b = \sqrt{\frac{121.78}{248.4}} \times 100 = 4.42\%$

$CV_o = \sqrt{\frac{90 + 121.78}{248.4}} \times 100 = 5.84$

$90 + 121.78 = 211.78$

$\frac{90}{211.78} \times 100 = 42\%\ V_w$

$\frac{121.78}{211.78} \times 100 = 57\%\ V_b$

$\sigma = \frac{\bar{R}}{d_2} = \frac{15.32}{2.059} = 7.44$

$CV_r = \frac{7.44}{248.4} \times 100 = 3.0\%$

SKEIN STRENGTH

Cone No.	Skein Strength X	$x - \bar{x}$	$(x - \bar{x})^2$	$x_i - \bar{x}_i$	$(x_i - \bar{x}_i)^2$	$\bar{x}_i - \bar{\bar{x}}$	$(\bar{x}_i - \bar{\bar{x}})^2$	R
1	248	0.4	.16	7.3	53.29			
	256	7.6	57.76	0.7	.49			
	261	12.6	158.76	5.7	32.49			
	256	7.6	57.76	0.7	.49			13
$\sum$	1021							
$\bar{x}_i$	255.3					6.9	47.61	
2	242	6.4	40.96	3.5	12.25			
	252	3.6	12.96	6.5	42.25			
	238	10.4	108.16	7.5	56.25			
	250	1.6	2.56	4.5	20.25			12
$\sum$	982							
$\bar{x}_i$	245.5					2.9	8.41	
	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.
24	239	9.4	88.36	3.8	14.44			
	250	1.6	2.56	7.2	51.84			
	242	6.4	40.96	0.8	.64			
	240	8.4	70.56	2.8	7.84			11
$\sum$	971							
$\bar{x}_i$	242.8					5.6	31.36	
25	249	0.6	.36	1.5	2.25			
	249	0.6	.36	1.5	2.25			
	240	8.4	70.56	7.5	56.25			
	252	3.6	12.96	4.5	20.25			12
$\sum$	990							
$\bar{x}_i$	247.5					0.9	0.81	
Sum	24844		20600.6		6750.0		3462.6	383
Avg. $\bar{\bar{x}}$	248.44							15.32

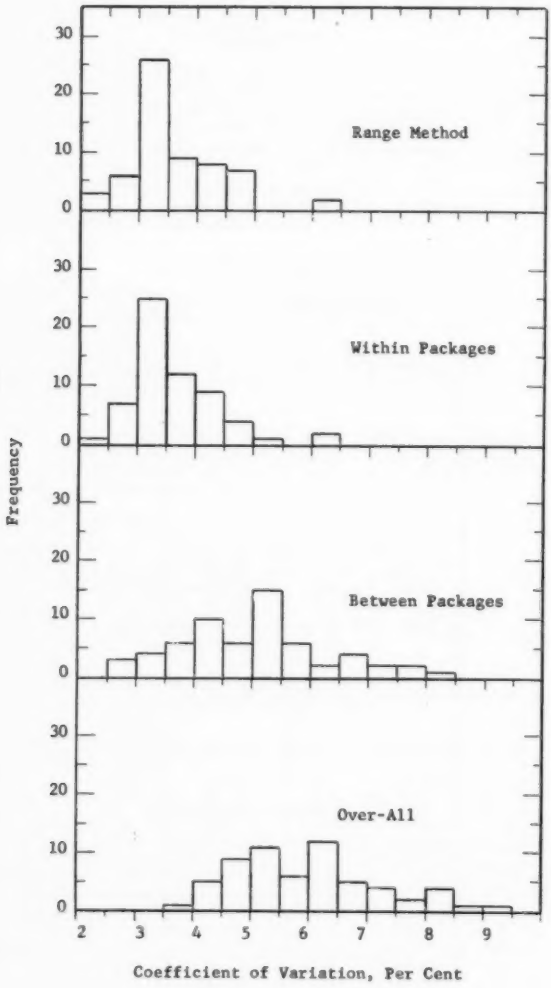


Figure 1. Coefficients of Variation for Skein Strength of Combed Cotton Yarns

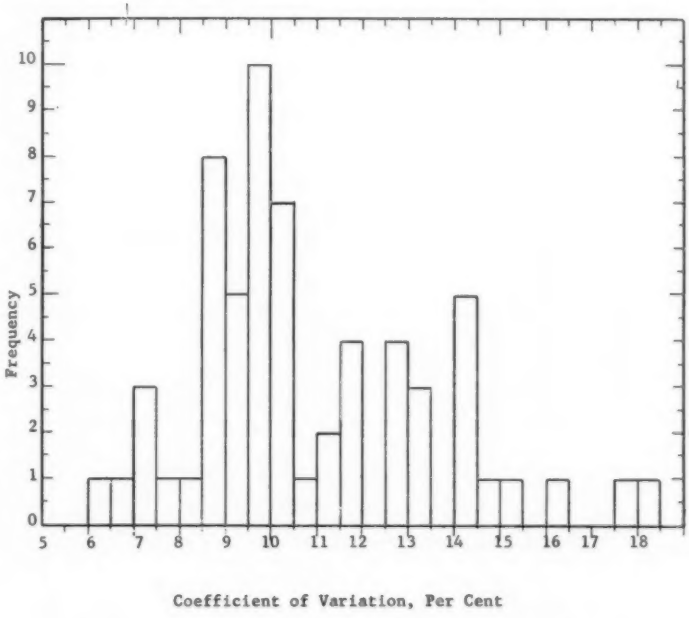


Figure 2. Coefficient of Variation for Single End Strength of Combed Cotton Yarns

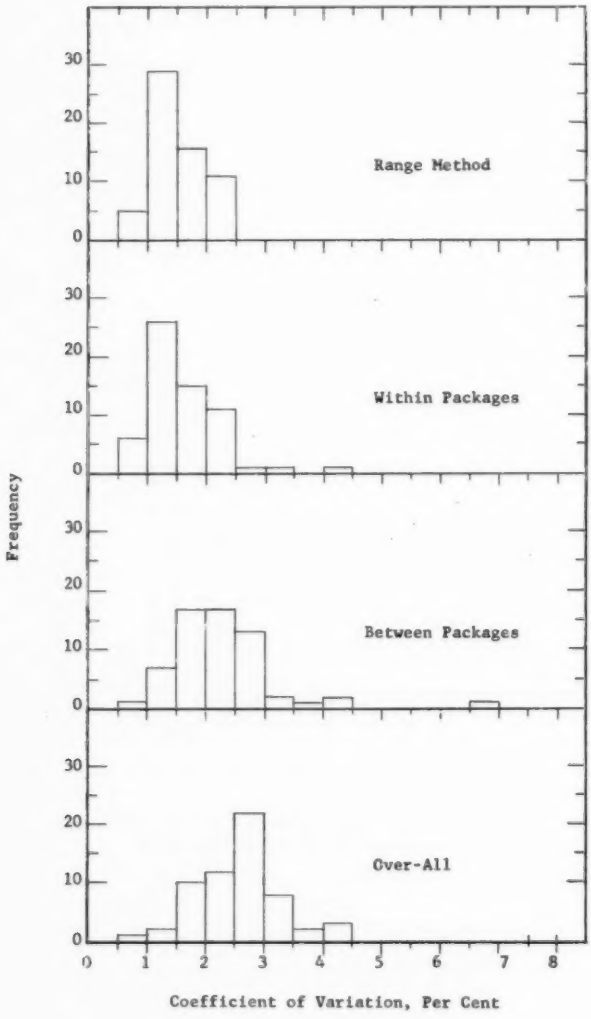


Figure 3. Coefficients of Variation for Yarn Number of Combed Cotton Yarns

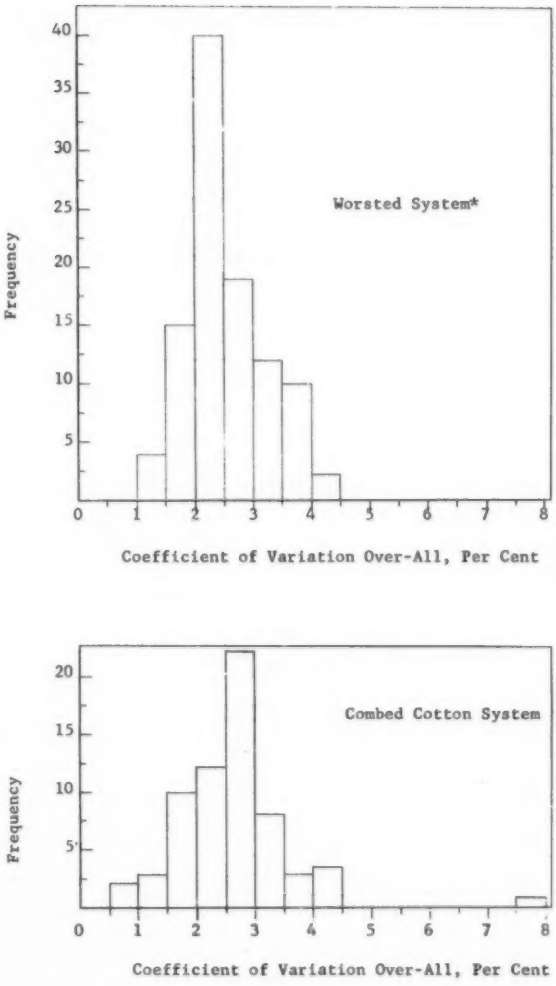


Figure 4. Comparison of Over-All Coefficient of Variation For Yarns Spun on the Cotton and Worsted Systems

* Data for Worsted System Yarn Courtesy of G. M. Bornet, Ontario Research Foundation, Toronto, Canada.

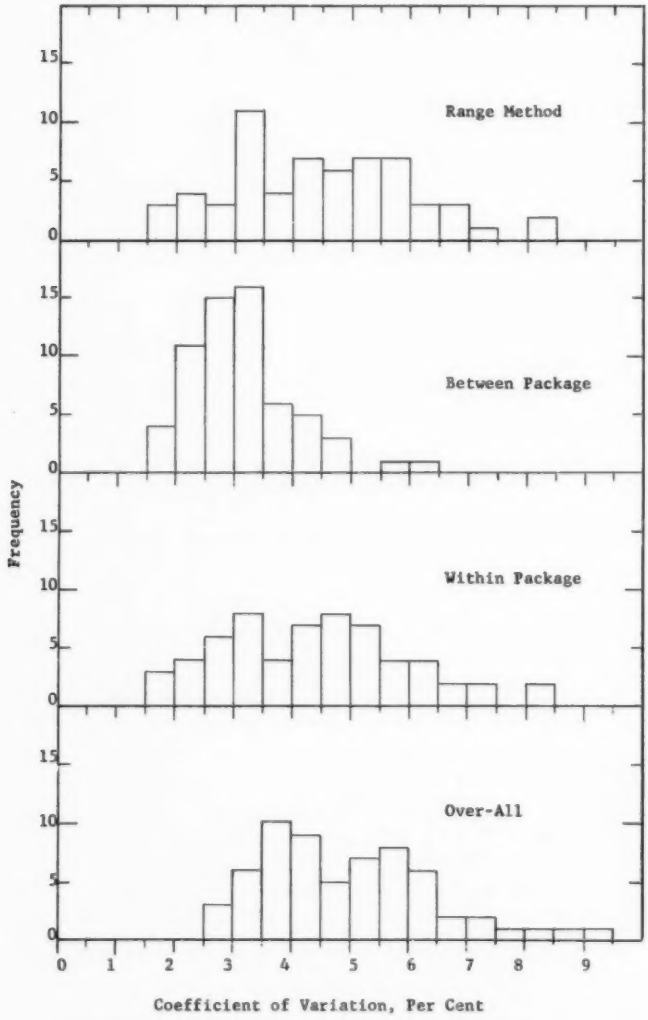


Figure 5. Distribution of Coefficients of Variation for Twist in Combed Cotton Yarns

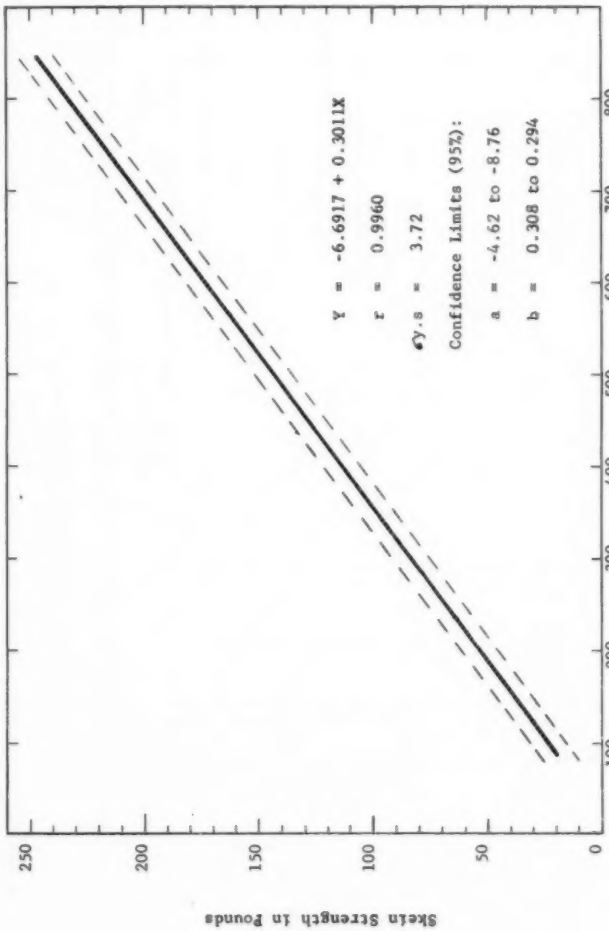


Figure 6. Relationship Between Skein Strength and Single End Strength of Combed Cotton Yarns

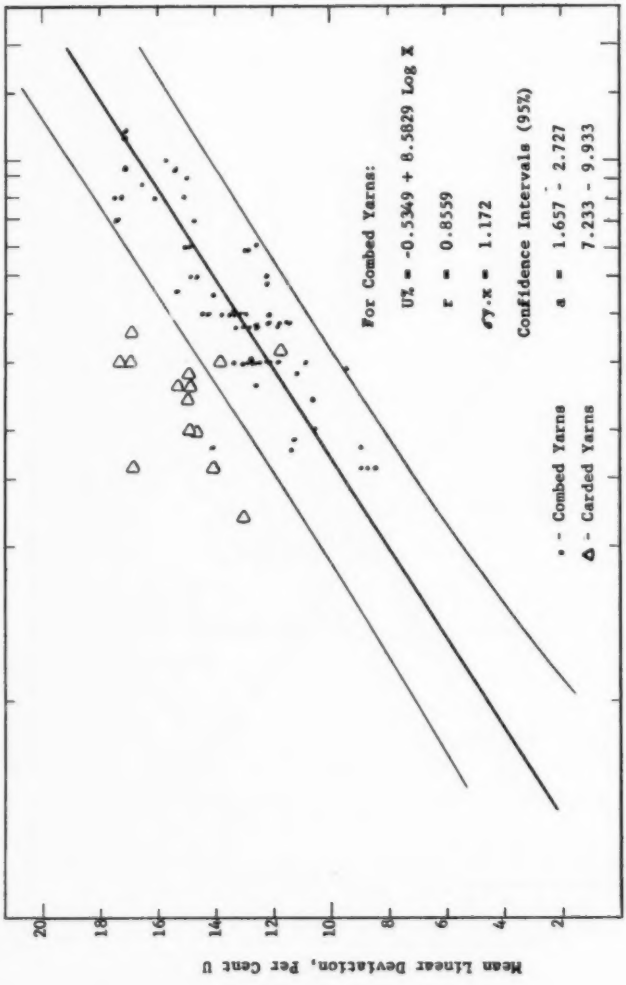


Figure 7. Relationship Between Yarn Number and Mean Linear Deviation as Measured by the Uster Integrator

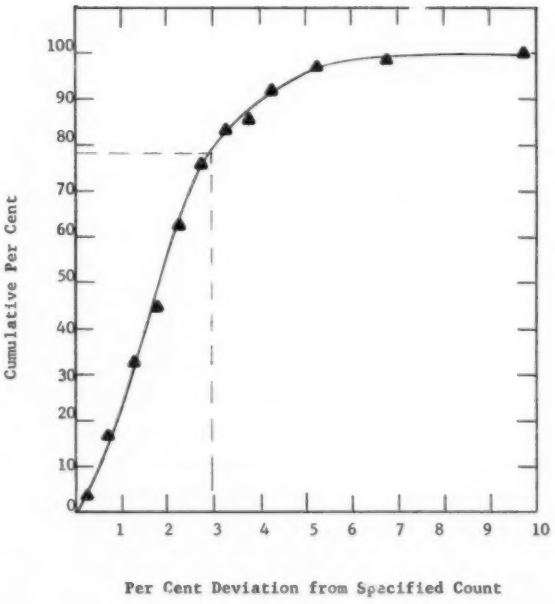


Figure 8. Cumulative Distribution for Per Cent Deviation of Yarn Number from Specified Count of Combed Cotton Yarns

Note: A tolerance of 3 per cent is specified by the American Society for Testing Materials

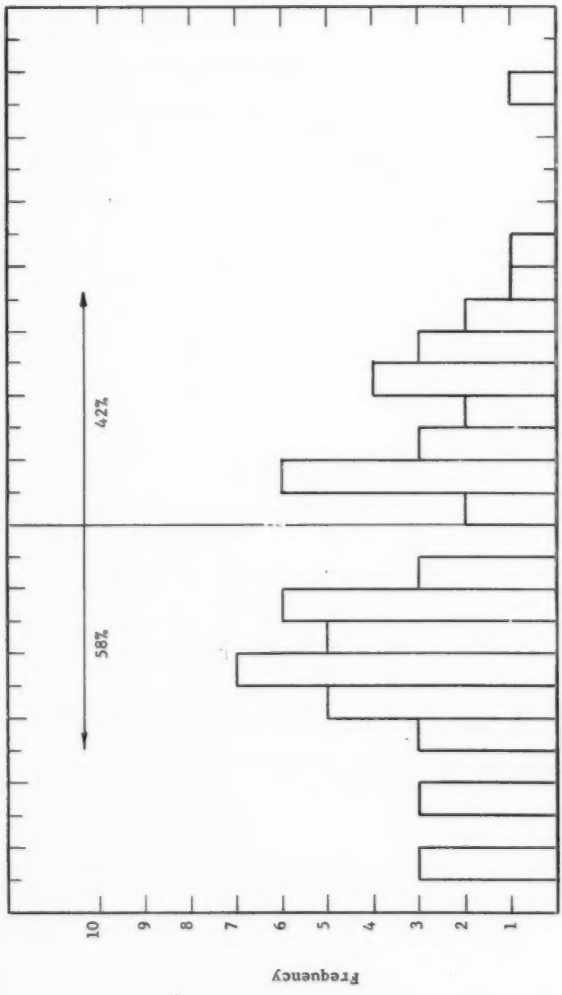


Figure 9. Per Cent Deviation of Actual Yarn Number From Specified Number for Combed Cotton Yarns

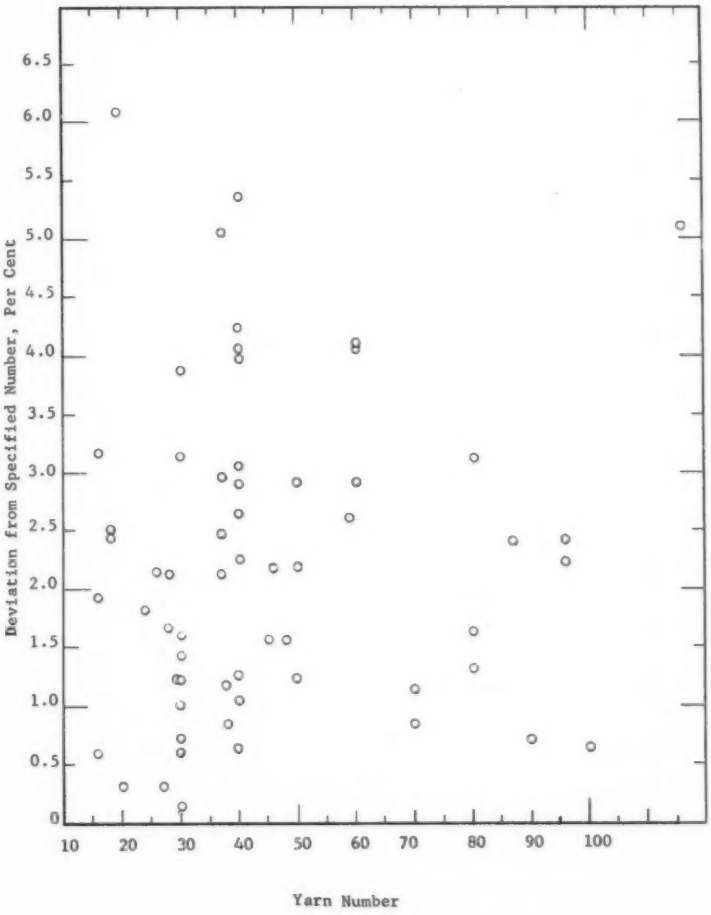


Figure 10. Relationship of Yarn Number Variation and Specified Number for Combed Cotton Yarns

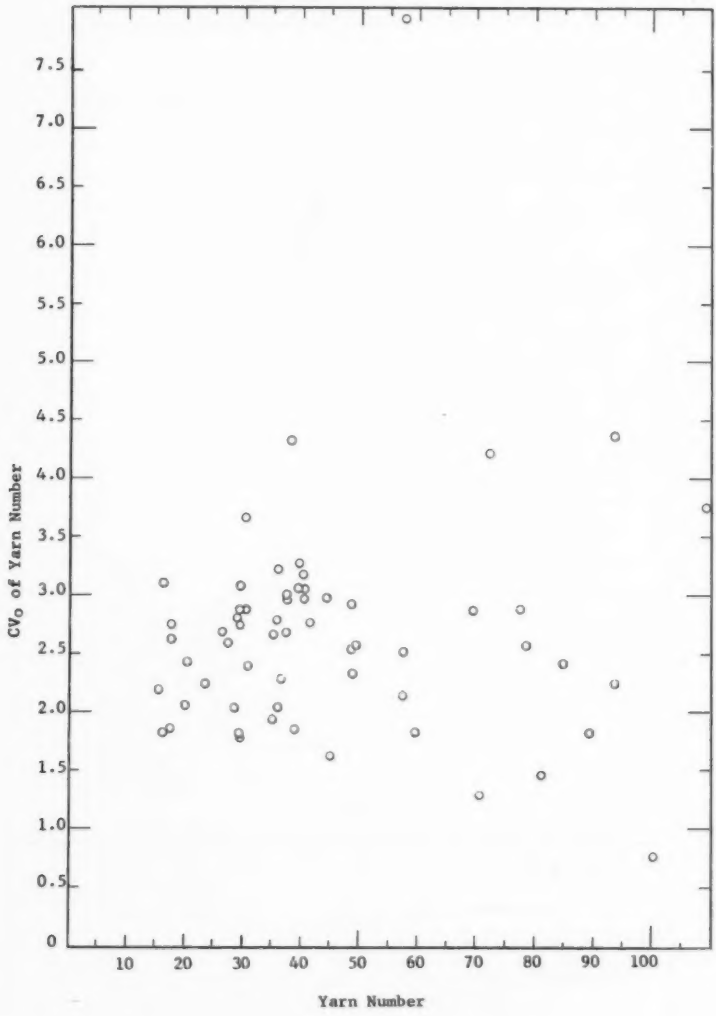


Figure 11. Relationship Between the Coefficient of Over-All Variation for Yarn Number and the Specified Yarn Number for Combed Cotton Yarns

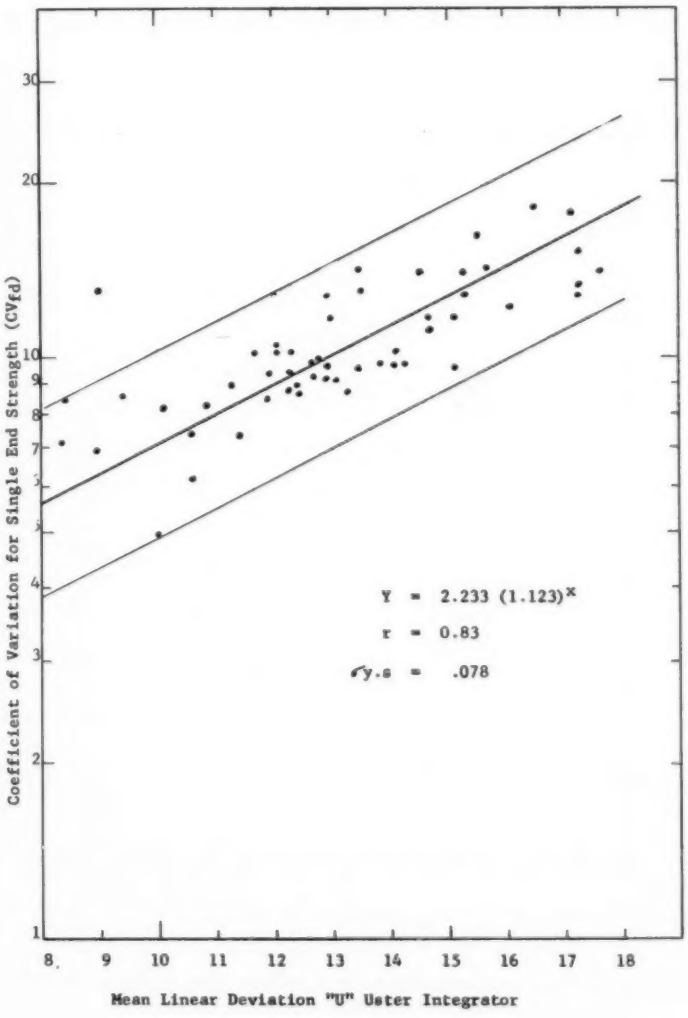


Figure 12. Relationship Between Per Cent Mean Linear Deviation and Coefficient of Variation for Single End Strength of Combed Cotton Yarn

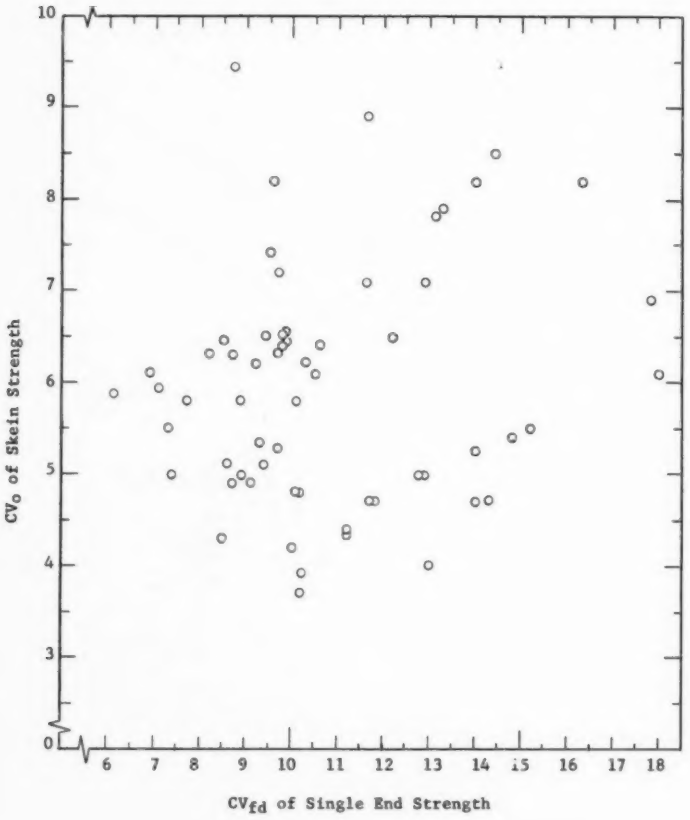


Figure 13. Relationship Between Coefficients of Variation for Skein and Single End Strengths

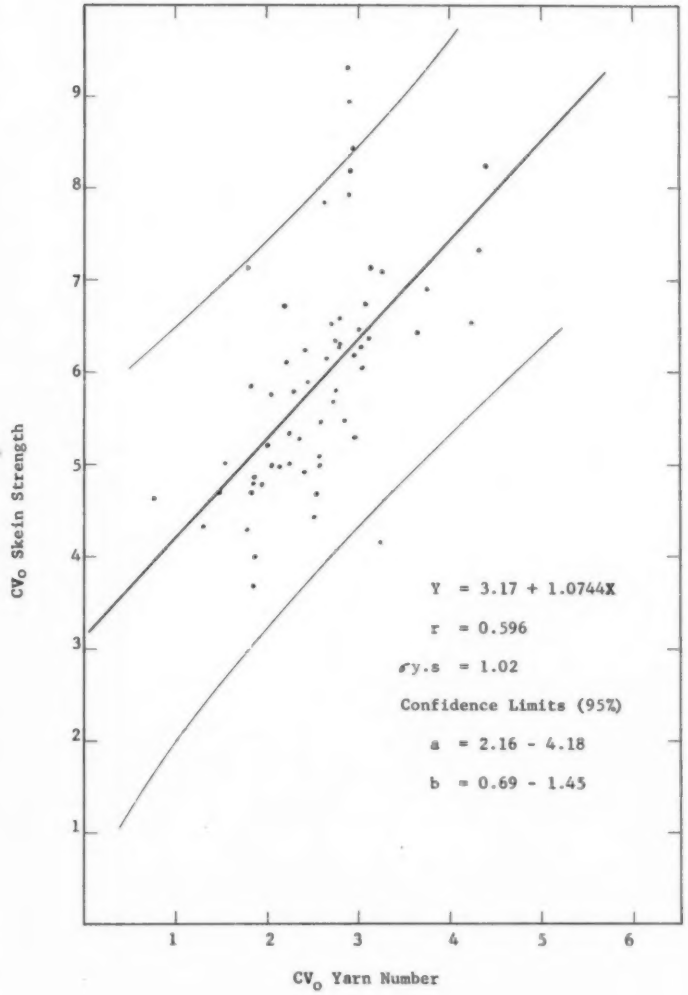


Figure 14. Relationship Between Over-All Coefficient of Variation for Skein Strength and Yarn Number

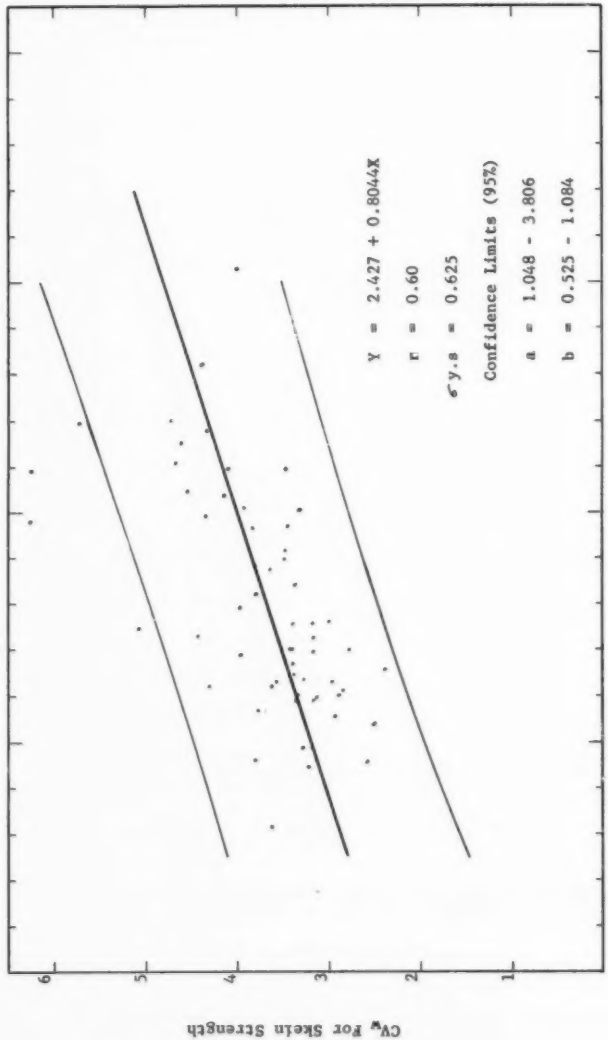


Figure 15. Relationship Between Coefficient of Variation Within For Skein Strength and Coefficient of Variation Within For Yarn Number

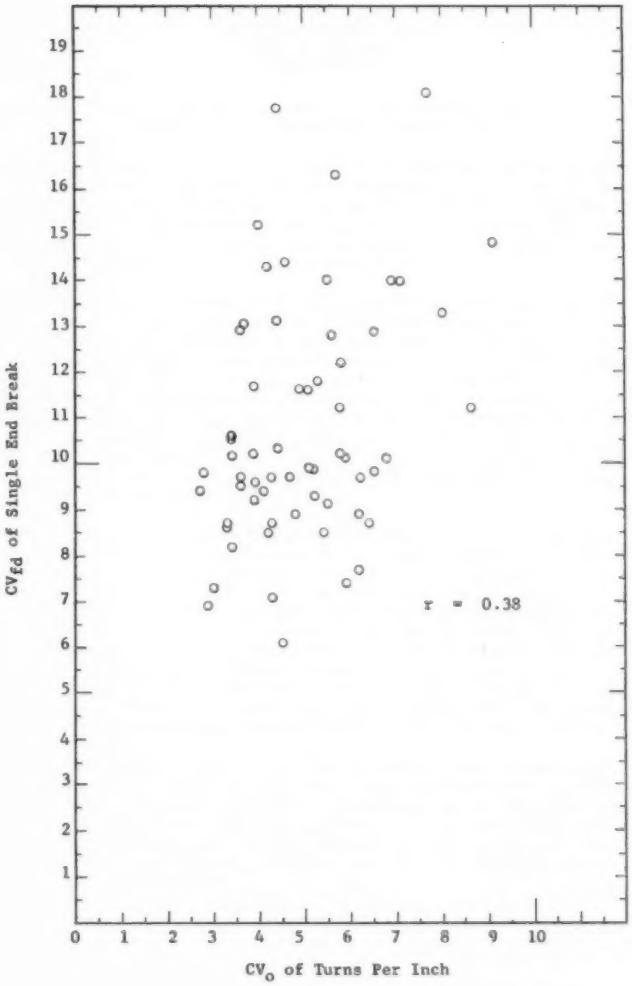


Figure 16. Relationship Between the Coefficients of Variation for Turns Per Inch and Single End Strength

SAMPLING PROBLEMS IN THE METALS INDUSTRY -INTRODUCTION-

A. L. Raich, Quality Control Statistician
Colorado Fuel & Iron Corp., Pueblo, Colo.

Sampling problems are as old as the era of mass-production techniques in industrial processes. When any individual product moves out of the category where perfect verification of its conformity to standards is feasible, either because of potential damage due to testing or because of production lot sizes so large as to make complete inspection unreasonable, some sort of sampling technique becomes necessary for quality assurance.

Since the metals industry has long been a basic producing source for almost all industry, it is reasonable to expect that those working in the metals industry will be in the forefront of practical application when new techniques of sampling are developed. When statistical sampling was still a relatively new concept, when the first government ordnance sampling tables were developed by Dodge and Romig, and during the still-current implementation of general attribute and variables sampling standards, metallurgical people were involved intimately with developmental work in this field. It is only fair to say that metallurgical voices have also been among the loudest heard in criticism of sampling standards they considered improperly applied, and in pointing out limitations of available sampling plans.

Consequently, the Metals Technical Committee of A.S.Q.C. has undertaken over the past several years to sponsor or co-sponsor Convention papers looking into problems of current and practical interest in the application of statistical sampling techniques. MIL-STD-105A, non-random defect patterns, and control of continuously extended products (last year's panel subject) are typical examples of subjects previously discussed. The papers which follow are an extension of these earlier discussions, taking up issues raised then and subsequently. For example, one problem of primary concern today is the producer-consumer relationship when a fabricator is contractually bound to a discrete sampling plan, but finds his primary supplier claiming that such plans are inapplicable to his continuously extended product sold as coils, rolls, or spools.

Represented on this panel are a primary producer and a fabricator from both ferrous and nonferrous metals industries. With this balanced representation, it is hoped that some generally applicable guideposts can be developed for cooperative efforts in working out sampling problems. It is, after all, the ultimate purpose of sampling techniques to make quality control procedures better and simpler, and not more difficult. Perhaps this type of informed discussion can further that worthy objective.

HOW SHALL WE APPLY IT?

L. J. Barker, Chief Specification Engineer
Kaiser Aluminum & Chemical Corporation, Oakland California

As members of the aluminum industry we have been asked to discuss the applicability of statistical sampling plans such as MIL-STD-105A to the products we fabricate for aircraft manufacturers. Our quality control procedures in general are more thorough and elaborate than statistical methods and are individually fashioned for each type of product. However, we recognize that statistical methods usually save money and, therefore, we would like to adopt them where we can. We continue to study all operations for possible application, but we do not see how they can be effectively utilized for control of most of the quality characteristics of aircraft items. Let us review some of the quality control methods we do employ and how it was proposed that MIL-STD-105A apply.

Government specifications for individual aluminum products--the material or commodity specifications--list four main quality requirements. The first two, chemical composition and mechanical properties, have long been "in control" and, in fact, a form of statistical sampling is maintained on a lot basis. Samples for chemical analysis are taken during the pour of 20,000 pounds of aluminum alloy. The melting operation including quantities, stirring, fluxing, etc. has become so refined through the years and so controlled that very little off-composition occurs. The relatively small sampling of the heat has been found sufficient to insure that the ingots cast are within specification limits. Frequency distribution data is kept of the quality of each element. This data is reviewed periodically, and adjustments in practices made as necessary. Similarly, samples are taken to determine the mechanical properties of the final product. Frequency distribution data is likewise kept and reviewed and adjustments in practices made as needed. Such quality control procedures are acceptable to both the aircraft industry and the government.

It was proposed that MIL-STD-105A apply to the other two quality characteristics--namely, dimensional tolerances and visual appearance. The aluminum industry has had no quarrel with MIL-STD-105A by itself since it is undoubtedly a well-eningeered document from a statistical quality control standpoint; however, there are serious problems connected with its applicability to the aluminum industry.

First, the reference to MIL-STD-105A in the material specifications required that "from each inspection lot a representative sample of plates or sheets shall be selected by the government inspector." To have the government inspector select the samples would require several more inspectors at each aluminum plant--at greatly increased cost to the government--and the selection of individual samples without damage to the others in an order would be extremely difficult at the proposed shipping stage.

Second, and more important, the aluminum industry does not feel that statistical sampling plans for these attributes are feasible within the conditions under which we operate. Ours is a high-speed job-shop operation.

Basically, there are four sets of complex variables which multiply to cause formidable stumbling blocks to the application of statistical quality control methods.

VARIABILITY OF PRODUCTS

The aluminum industry produces a tremendous variety of products. As an example, our sheet and plate rolling mill at Trentwood, Washington, produces for aircraft consumers alone the following variety of products, the specifics of which must be spelled out on purchase orders:

Alloys	-- 5 (to name only the most popular)
Bare or Clad Surfaces	-- 2 (actually several different cladding alloys are offered)
Tempers	-- 3 (O, T3, T6--more available)
Finish	-- 2 (mill or skin)
Stress Relieved or Not	-- 2
Special Ultrasonic Quality or Not	-- 2
Total Multiple	-- 240

This total of 240 varieties or combinations of variables does not include the almost limitless combination of sizes--various thicknesses from .006 inches to 3.000 inches; widths from, say, 1 inch to 100 inches; and lengths from, say, 1 foot to 20 feet and more. And so far we are counting only aircraft items. When you multiply this by the greater tonnage of commercial business all produced on the same pieces of equipment, the possible combinations in one sheet and plate mill are infinite. By grouping items or variables together as shown here, including commercial material and the more common standard sizes, our accounting people have come up with a figure of 60,000 possible individual sheet and plate items or products--any of which we might be producing at any given day.

Photograph showing some of the various items produced in a sheet and plate plant.

Naturally, these products are not all run on one piece of equipment; however, all are run through the first hot mill, for example, and rolled to various predetermined dimensions at different temperatures for different alloys and with consequent different degrees of ease. Some 300 ingots are hot rolled each day; that is approximately three minutes per ingot. Hot line lots may vary from one ingot produced to one set of predetermined dimensions, etc. to 20 or more to the same dimensions.

After hot rolling, a multitude of operations in many combinations are possible--cold rolling and annealing cycles in one to several stages, heat treatment, aging, cold rolling on two-high mills, slitting, shearing, sawing, stretching, etc.

With such variation in product it is evident that a given piece of equipment may only be operating a few minutes on one setting for one lot or a group of similar lots and then it is reset for another set of conditions. This

occurs in spite of the fact that our planning and scheduling people take maximum advantage of grouping lots so a minimum of resetting is necessary.

VARIABILITY OF LOT SIZE

With such a variety of products from heavy plate to thin sheet it is only natural that the lot size or minimum number of similar units per lot would have considerable variation. Final product quantities actually vary from 1 piece to 1,000 pieces. In the sheet and plate plant we average over 100 such lots per day totalling about 1,000,000 pounds.

PRODUCTION VARIABLES WITHIN A LOT

Coupled with the fact that we do not produce a constant product, each product is not uniform in itself. Nor are the inconsistencies of uniform distribution. An example of this would be special tolerance sheet.

The in-process material received at the two-high or so-called skin mill for the final rolling is a stack of flat sheets continuously cut from a coil with a nominal thickness of about .002 inches greater than the final thickness ordered. The thickness is not uniform from sheet to sheet or even within any one sheet. This non-uniformity is caused by variations in the gap between rolls of the previous four-high mill on which it was rolled--due to the heating and expansion of the rolls, possible adjustments, and a non-parallel gap between rolls caused by the "crown" of the rolls.

These non-uniform sheets are fed one at a time into the two-high mill where in one pass each sheet is reduced to final thickness. Compensations must be made for previous non-uniformities. Additional variations are caused by the heating and expansion of these rolls and possible adjustments. The final lot is sheared to width and end-sheared to length.

Special thickness tolerances in a lot of metal is a quality then that cannot be measured by random in-process sampling. The thickness changes with time but not as a linear function since adjustments are made in increments. Quality control demands that frequent measurements be made to see that the close tolerances are being met. Experience on how often to measure and where to measure is of vital importance.

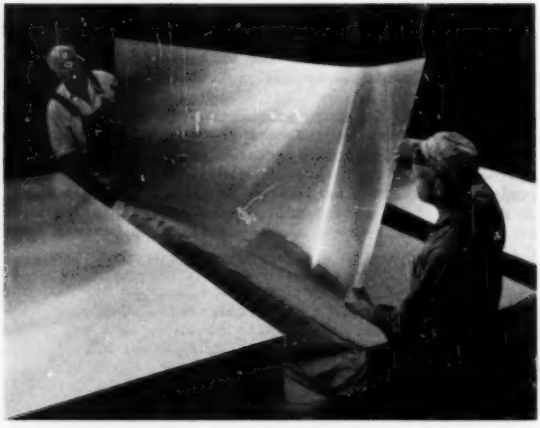
An important point here in addition is to realize that the two-high mill may have been idle or may have been rolling a different thickness previously and after this operation of a few minutes duration will be reset for a different thickness. This is typical of all the equipment in the plant--none of it runs on a continuous basis for just one product.

Other examples of variable quality within a lot are defects such as oil stain, heat treat stain, blisters, scratches, etc. Any one defect may show high incidence in one month and none the next, etc. They are truly of a non-constant nature. Such defective material is weeded out as it occurs and the fact reported to the operating departments for corrective measures.

VARIABILITY OF INSPECTION

Because of the former variables--variable product, variable lot size and variable quality--it follows that inspection techniques will have considerable variation. We do not attempt to make one quality. Rather, we attempt to build into the product those characteristics the customer requires. For example, the same aircraft manufacturer may place a dozen orders for sheet or plate materials each with a different set of production and inspection requirements. One order may require 100% in-process ultrasonic inspection to a high standard. Here special practices for casting ingots are employed and, in effect, those ingots that pass the ultrasonic inspection are the cream of the crop. Another order may require half-commercial tolerances which are non-standard requiring additional control and inspection. Another may be skin quality meaning a brighter than usual surface as free from blemishes as possible. A third order may be specially stretched for stress relief qualities with each plate marked with bench marks..., etc.

Inspection becomes a matter of deciding whether or not we have met the customer's requirements. Techniques for one lot of material, for example, may require that every piece be examined carefully for surface defects since virtually none can be tolerated for the end use the customer has in mind. Another lot of similarly produced material may require only that normal precautions be taken and surface quality is of no concern. Another order may require that dimensions be extremely accurate; so measurement of every piece may be necessary to insure conformance.



Photograph showing combined methods of final inspection for surface blemishes and packing for shipment.

The inherent variability of some of our operations and machines is so great compared with certain of our customers' requirements that random sampling in such cases would result in rejection of all lots. The only way to meet many customer requirements is to hand select individual pieces by 100 percent inspection. If possible, the remainder of such lots is applied to an end use with less stringent requirements.

Each aircraft item, in effect, is a tailor-made product, produced in a tailor-made fashion and inspected in a tailor-made manner.

SAMPLING PROBLEMS IN THE METAL INDUSTRIES
(Non Ferrous Metals)

J. Y. McClure, Manager of Quality Control
Convair, A Division of General Dynamics

How much sampling is enough; this is the question?

Too often when considering sampling and sampling plans, there is a tendency to put the cart before the horse. We pick a plan that we like or one that someone else has recommended highly, and then attempt to adapt our processes to fit this plan. Occasionally this works, but more often than not it fails. Why does it fail? The simple reason is this, in most cases the process is not flexible enough to be altered to suit the sampling plan. For some reason or other, people fail to realize that the flexibility required is in the establishment of the sampling plan and not in the process to be sampled. The proof of this is the many good plans already in existence. Some of these can be used just as they are, because the situations under which they were developed are similar to the one we wish to use them on. Others, however, although they have no ready application, can serve as useful guides to the building of applicable sampling plans.

How much is enough? This is where one must start when deciding on whether he can go to sampling inspection. How many pieces must be inspected, or how many analysis must be taken to assure a needed quality level is present, within limits of course.

It is a common practice to establish limits which are too severe. We do this thinking we are achieving a greater degree of protection. This is literally true, yet the added cost may not be justified, the degree of protection beyond that which is needed is often costly excess baggage.

The total picture must be examined when setting up sampling plans. Will the defects overlooked in the involved risk of sampling be caught some place else? If so, will the cost of finding them simply be that of laying aside a misfit or will it require expensive teardowns of complex assemblies?

"Too much" can be wasted expense, by the same token "too little" can cause expensive cost in later stages of production, sampling must do a specific job. For it to do this job, we must be cognizant of that which needs to be sampled. Is it the variation of a machine, the efficiency of a group of workers, the differences within a lot of bulk materials, or even a company's integrity? Once this is determined a plan can be adopted to fit the process.

In our metals receiving area, at Convair - Fort Worth, all these things were taken into consideration, and we finally decided that what we would sample would be integrity. In other words could we rely on our vendors' test results. To do this required that we accomplish two basic tasks. They are as follows:

1. Establish reliability of the vendor
2. Establish reliability of his product

Establishing vendor reliability begins with an initial vendor evaluation survey made by representatives from the quality control department. The purpose being to establish vendor capability as far as product and test facilities are concerned. Once this has been accomplished and the vendor is rated as satisfactory the initial order can be placed.

All receipts which follow during the next three to six months are sampled and analyzed to see if the vendor's test results compare with our test results, or in the cases where the vendor certified that certain minimum specifications have been met, we will test to validate this certification.

When the vendor and product reliability is established, we can, with confidence, depend on vendor's test results or certification. This confidence is maintained through the use of a random sampling plan, which calls for periodic sampling of lots received from each vendor. Should one of these samples fail to meet specification, the vendor is notified and a minimum of 5 consecutive lots are then checked to be sure corrective action has been taken.

This outlines the general approach that we are using. To be sure, there are some items which require 100% chemical and physical checks, where as, there are others that receive no more than simple dimensional and hardness checks. Here again those factors discussed earlier have been taken into consideration and an optimum overall program has been accomplished.

STATISTICAL EVALUATION

Paul S. Olmstead
Bell Telephone Laboratories, Inc.
Whippany, New Jersey

INTRODUCTION

In recent years, attention has been directed to statistical estimation, statistical design of experiment, analysis of variance, response surfaces, and random balance. In each case, mathematical theory has assumed that all residual variation has been random. When the residual variation is small, making this assumption is often considered to be unimportant. As a result, as statistical techniques have been applied by statistician-engineer teams, randomness has come to be accepted as a characteristic of residual variation. Orinarily, no test is made to see whether this assumption really applies.

Basically, the quality control chart was devised as a means for detecting statistically the presence of assignable causes that produce nonrandomness in residual variation. In particular, it emphasized the importance of associating "other information" with points found outside control limits. It is quite clear that this principle should be introduced into the analysis of all statistical data regardless of the source. Otherwise, many operationally important conclusions of statistically designed experiments may be overlooked. In other words, the plan for such an experiment should also include provision for examination of residuals for control and determination of the probable presence of important assignable causes. This is apt to lead to statistical evaluation of experimental data differing considerably from customary analysis.

Clearly, statistical evaluation must take into account the purpose of the evaluation. For a long time, statistical consideration was limited to presenting statistics as history. Obviously, this relates to Statistics Past. Evaluation is also of interest in connection with Statistics Present and Statistics Future. It is here that Quality Control became important, first as a basis for discovery and then for prediction.

STATISTICS PAST - HISTORY

The earliest use of statistics was to record what had happened. In fact, the original constitution (1)* of the American Statistical Association (formed November 27, 1839) stated that: "The objects of the Society shall be to collect, preserve and diffuse Statistical information in the different departments of human knowledge." At that time, the decennial Census was perhaps the most important interest of statisticians.

A large part of the data obtained for historical purposes was presented in tabular form. This is often true even today. However, more limited summaries of the data are sometimes sufficient. Although most of the desired information about a distribution may be related to its average, the distribution is more clearly defined by also knowing its standard deviation, skewness, and kurtosis. Karl Pearson (2) systematized the available information about frequency distributions by deriving a generalized frequency curve based on a single differential equation, with parameters obtained by the method of moments. Thus, he made it possible to identify a distribution by its first two moments (average and variance) and a distribution type roughly defined by the third and fourth moments.

*See References

Pearson recognized that experimental data were usually insufficient to provide even approximate estimates for the higher moments. In fact, even the third and fourth moments may be insufficient to distinguish among several distribution types. For instance, the normal law is a limiting case for seven of the Pearson type distributions. In view of this, selection of an assumption regarding distribution type is usually made on theoretical grounds. Experimental determination of distribution type is quite hazardous. In part, this is because conditions under which data are obtained are often far from controlled. In addition, a reasonable estimate of skewness requires approximately 1000 observations and 10,000 may be required for a good measure of kurtosis. As a result, the most informative items relating to historical statistics are often taken to be the average, the standard deviation, and theoretical grounds for a distribution type.

In some cases, there is interest in certain correlations that have been observed. This has been a very controversial subject. The regression of Y on X may be real but the correlation coefficient calculated to illustrate it may have no meaning at all. There have been many instances published giving cases of nonsense correlations. Such spurious correlations are sometimes obtained simply because each value of Y is strongly dependent on its predecessor in the ordered sequence in which it was obtained, and a similar statement is also true of each vlue of X. Thus, neither the Y's nor the X's are random samples of a simple constant cause system. In this case, there may be a meaningful regression of Y on X related to the dependence but not a meaningful correlation. A more complete historical picture as it relates to a meaningful correlation then might be control chart presentation of the ordered sequences of Y and X data separately, the regression of Y on X, and the value of the correlation coefficient (if the control charts for both X and Y show control). The purpose of this limited discussion of correlation is to suggest that understanding the process that produced it is more important than the correlation coefficient itself.

STATISTICS PRESENT - DISCOVERY

The use of statistics for purposes of discovery and identification of an assignable cause is quite recent. When Shewhart (3) set limits on his control charts and found that he could usually find a reason for each point that fell outside these limits, he was using statistics for discovering assignable causes. Tying action to each discovery means that statistics present is involved. In many applications of quality control, the charts are placed at the machines so that the action need not be delayed. This avoids the dangers associated with the use of statistics past for this purpose. Conditions may have changed from what they were when the data were taken. Thus, new statistics may have to be obtained to be reasonably certain that action suggested by statistics past is still appropriate.

In manufacture, some of the possible causes of out-of-limit data are known prior to start of production, i.e., before data are obtained that suggest their presence. If only one cause could be responsible for the observed type of outage, the corrective measure is relatively easy to determine. If more than one cause could be responsible for the outage, additional investigation may be required to determine the proper action.

The major cue to making discoveries with statistics is subgrouping of data. There are many possible sources of subgrouping.

Some of these arise from the design of the experiment; some from the time sequence in which the experiments were made; others arise from collateral data, often called "other information", that may have been recorded during the experiment. There are, of course, occasions where statistical evidence must be used as a reason for discussion and attempt to recall "other information" that has not been recorded but of which the experimenter is aware. Such recall is most effective when the analysis is made while the experiment is in progress. A long delay between the time of the experiment and the analysis may postpone an important discovery simply because relevant "other information" has been forgotten.

STATISTICS FUTURE - PREDICTION

As indicated in the introduction, statistics are useful not only for summarizing past data and for making discoveries but also for purposes of prediction. In establishing a sound basis for statistical estimation, R. A. Fisher(4) devised a method for identifying the statistic that would be most "efficient" for estimating a selected distribution parameter from a random sample. The major restriction on the use of such results in a practical example is that observed data may not be a random sample from the assumed universe.

Shewhart recognized this in connection with his work on statistical quality control. He also recognized that "Student ranges"(5), i.e., confidence intervals, associated with parameter estimation were apt to be confused with tolerance ranges or limits as normally used in engineering. He showed that reliable estimates of tolerance limits to include a fixed fraction of a "universe" could be obtained only from large samples, whereas the probability that the limits associated with "Student ranges" would include the average of the "universe" was independent of sample size. However, a recent article(12) suggests that decisions based on tests of significance become more informative when related to confidence intervals determined for larger sample sizes.

That nonrandom variation is apt to exist in a series of consecutive measurements has long been accepted by engineers. It was not clear that something could be done about it, however, until Shewhart devised the quality control chart. This assisted in the discovery, identification, and elimination of assignable causes to the point where a production process would provide an output that would meet criteria for statistical control. Under these conditions it was possible to make meaningful predictions concerning the quality of the future output.

As Shewhart has pointed out(5), there are three factors associated with reaching such a conclusion:

1. The use of statistical theory to establish for a given universe the probability of meeting a given criterion,

2. Knowledge that appropriate action has been taken to eliminate assignable causes that have been discovered and that this has been followed by a sufficient quantity of data meeting the criterion of control, and

3. A belief based on experience that having met the criterion of control in this way it is safe to assume that the state of statistical control that it indicates can be maintained.

It is quite obvious that there must a one-to-one connection between the statistical model and the attained state of statistical control for predictions to have maximum validity. Evidence that this connection exists would be provided in the simplest case by a

control chart of twenty-five consecutive averages of four that meet the criterion of statistical control(5). However, it must be followed by a continuing operation of statistical control designed to be self-correcting for the purpose of attaining the desired state of control.

Engineers are also interested in making good predictions in cases where a state of statistical control has not been realized. In fact, this is the usual case in most research and development work. In the simple case, where interest is centered primarily on a single variable, good prediction depends on the kind of statistical control that is inferred from previous experience, the assignable causes that have been found, their probable effects, and knowledge about the steps that have been and will be taken to assure uniformity in the quality characteristic with respect to its desired level.

Thus, good prediction depends in part on statistical information, in part on an evaluation of this information, and in part on knowledge about the system that will be the source of the future statistical data. Statistical evaluation includes all statistical statements relating to prediction that apply to specific assumed systems and their probable statistical outputs. Statistical evaluation is based on the evidence provided by past data that would be applicable to the conditions under which future data will be obtained.

STATISTICAL EVALUATION - SPECIFIC CASES

As noted above, prediction is most meaningful when it is supported by evidence of statistical control. For example, if at least twenty-five averages and ranges of consecutive samples of size 4 have been within control limits, and this situation was reached by an operation of statistical control that resulted in detecting and eliminating assignable causes of variation, then a statement of statistical evaluation such as the following would apply:

STATISTICAL EVALUATION WITH EVIDENCE OF STATISTICAL CONTROL

Assuming a continuing operation of statistical control, this process will continue in a state of control within limits based on the average $\overline{X}$ and the average R determined from the 25 averages and ranges of the consecutive samples of size 4. This is a statement based on experience that has been illustrated many times in Quality Control work.

In research and development work, the condition that 25 consecutive samples provide evidence of control is seldom met. The problem, then, becomes one of salvaging information. A somewhat common case is where the 25 averages include at least one outside control limits but the 25 ranges are all within. On the assumption that the cause of the outage has been determined and that it can be eliminated in future tests, it is possible to make a statement of statistical evaluation such as the following:

STATISTICAL EVALUATION WITH VARIABILITY ONLY CONTROLLED

Assuming a continuing operation of statistical control and elimination of discovered assignable causes, this process is capable of being established in a state of control within limits based on an average X to be determined and the average R determined from the 25 ranges of the consecutive samples of 4.

A tentative estimate of the average X that is relatively free from the bias introduced by a single outage is the median of the 100

measurements. (In some cases, the median of the 25 averages may be adequate.) It is frequently true that ranges will be found in control even when the averages are not. Knowledge that the variability is controlled is a big help in running down the cause of an outage in average. The reason for suggesting the use of the median as an estimate of average X in place of the average of the X's remaining after the outages have been omitted is that the direction of the outage may be important to establishing the average even though the magnitude is not. A second reason is that observations tend to cluster about the average so that a median obtained without eliminating the out-of-limit case or cases is still a good estimate of what the average would be if the causes of the outages were identified and eliminated.

Where more than one outage occurs for average X, the locations of the outages and their directions often provide information to help locate the cause. Some examples are:

1. Outages close together and in the same direction often indicate a shift in the cause process related to a change at a specific point.

2. If the first outages are in one direction and the last in the opposite direction, there is an indication of a trend that might be associated with measurement.

Obviously, each of these provides a clue[6] to the type of assignable cause that may have produced the observed effect. The engineer is usually aware of various physical causes, each of which could produce an isolated shift in level, a permanent shift in level, a periodic shift in level, a change in variability, and the like. By identifying the statistical change, it is therefore easier to associate the specific physical change responsible.

By providing this assistance to the engineer, his major interest may then relate to how well his process can be controlled after he has found and eliminated his assignable cause. In the statement indicated above, it is assumed that he is able to make adjustments to bring his average to the level that he desires. Since his individual ranges stayed within control limits, it is reasonable to believe that a statistically controlled process is possible to achieve.

Another situation worthy of consideration is where one or two of the ranges are out of control limits, but all averages are within. Ranges out of limits are not too common and it is seldom the case that more than two out of 25 would be outside. In fact, if a range is outside, the average of its sample is also likely to be outside. In dealing with samples of four, the lower limit for range is zero, so the only possible direction for the outage is to be too large. This coupled with knowledge gained from experience suggests that an individual X may be the cause of the outage. The individual X that is most suspect is thus the one that is farthest from the average of the group. If this is also the largest of all 100 measurements and the next largest is considerably smaller, a recheck of the measurement may show that the original X had not been properly recorded. Sometimes, such a recheck is not possible and the doubt raised by the outage cannot be removed by physical check. In this case, an alternate method of estimating control limits may be tried. As an example, the median range of the 25 ranges may be used to obtain the limits for the control charts. The various control factors based on the median range were listed by Ferrell[7]. With a strong possibility of at least one discrepant observation, it may be better to use the median of all 100 measurements as the central value for setting control limits for averages, using the factors

obtained from the median R. In this case, it is possible to make the following statement of statistical evaluation:

STATISTICAL EVALUATION WITH AVERAGE ONLY CONTROLLED

Assuming a continuing operation of statistical control and elimination of discovered assignable causes, this process is capable of being established in a state of control within limits based on the median X of all 100 measurements and the median R of the 25 ranges of the consecutive samples of 4.

Statistically, this statement is less desirable than either of the two preceding. However, it is much to be preferred to either of the others when applied to this particular case. Among the reasons for this are:

1. The median of the 100 measurements is less influenced by the individual X found responsible for the out-of-control range (or the two individual X's, if two are out).

2. The median R is less influenced by out-of-limit ranges which would have a large effect on the average R.

The use of median X and the median R to obtain control limits when all averages and ranges are in control is less desirable than using average X and average R because they are subject to large possible variation. However, where interest centers in determining the capability of a process that is in process of shakedown, the contribution of this factor is often considered to be of less importance than having a measure based on all the data. In this sense, the above statement represents an interim statement that may be made now but would be amended when statistical control is established.

STATISTICAL EVALUATION - GENERAL

These three simple cases show some of the way in which statistical data may be evaluated for use in making predictive statements. In the first case, where statistical control had been demonstrated, the usual rules of mathematical statistics are applicable. In the other two, experience in dealing with problems of this type must be taken into account. Experience, in this sense, can be summarized, at least in part, as follows:

1. Statistical test is useful in identifying causes by their possible effects on data, such as:

1.1 A maverick, gross error or blunder (shift in an individual),

1.2 Shift in average or level,

1.3 Shift in spread or variability,

1.4 A trend or long term cycle (gradual change in average or level),

1.5 A short term cycle (regular pattern of change in level), etc.

Experience has shown that when such effects have been demonstrated the engineer has usually been able to find a correctable cause(6).

2. A repetitive process that has not been examined for control by statistical methods and that has not been brought into control is very unlikely to be in statistical control(8). In other words, research and development data are almost certain to show at least one of the types of nonrandom variation mentioned above. Statisticians are still looking for a case where control exists before applying statistical methods.

3. Useful information relating to variability and often to average level may be frequently salvaged by following methods of statistical evaluation that take into account such factors of experience as the following:

   3.1 Most research and development data show sporadic periods of control that are typical of what the process might be after assignable causes have been corrected.

   3.2 The measure of variability (e.g., the range in the example cited above) is often found to be in statistical control even when the average is not. The estimate of variability obtained under these conditions, however, will usually be found later to be too high but until the primary trouble associated with the average is located and corrected, a better estimate, i.e., smaller variability, can impede its identification by drawing attention simultaneously to both primary and secondary effects.

   3.3 When the measure of variability is not in control, a single measurement is apt to be the culprit. Since error in recording is a very frequent cause of such a measurement, particularly in the cases that have been verified, the use of an alternate method of estimating variability as applied to the process is usually justified.

In this presentation, only those factors in statistical evaluation that are related to control charts have been considered. However, the ideas developed here are also applicable to analysis of variance situations where attention must be paid to residual variance to verify the assumption of randomness. Such verification is necessary to provide assurance that this basic assumption has been met. This was pointed out by Terry(9). However, it is common practice among statisticians to assume that such randomness exists, even though the means of testing it are so readily available. Unfortunately, if this test were made, many analysis of variance results would be found to be subject to error.

It is clear that the ideas related to statistical evaluation should be applied to all situations where statistical findings are to be used for the purpose of prediction. At the same time, it must be admitted that mathematical statistics has not as yet provided "exact" applicable tests. However, there have been at least two articles that have partially bridged this gap(8) and (10).

Each of these indicates that the penalty paid for superimposing additional requirements on a simple test is small. The reason that this is so is that the additional requirements are to detect the same type of trouble. As a result, failure to meet the additional requirements is not independent of failure to meet the simple test. This suggests that useful tests relating to statistical evaluation can be devised and the appropriate probabilities for the null hypothesis established. One reason for wishing to determine this is to obtain a measure of the risk of looking for a nonexistent cause.

As a first step toward the solution of this problem, it seems reasonable that use should be made of a high speed computer to generate a large number of random series of 100, each of which would be analyzed by the usual Quality Control procedure. The fraction of these series failing to meet the composite criteria for control would be determined and each nonconforming series identified by test failed. Thus, if the composite criteria applied to averages and ranges of the 25 samples of four, it would be possible to separate the series into four types:

1. Those meeting both criteria
2. Those nonconforming for average only
3. Those nonconforming for range only
4. Those nonconforming for both average and range

Each series, of course, would generate its own set of criteria which would be used to make the above classification. To find out how well such a set of criteria would apply as a prediction of future performance, it is necessary to determine the fraction of similar sequences of 100 that would meet the same criteria. Preferably a separate set of sequences should be used for this purpose. Such a simulation study would be expected to provide:

1. An estimate of the probability that a random series of 100 will meet criteria (no standard given) for control for averages and ranges of fours as applied in Quality Control.
2. An estimate of the probability that future series of 100 from the same universe will meet criteria established by 1. above.
3. A means for investigating the effect of using alternate ways of establishing criteria to be met in future series of 100 from the same universe when the test series does not meet the usual criteria for statistical control for a specific type of outage.

So far, attention has been directed to the simple cases only, where an eventual normal law type distribution is assumed. Obviously, the problems multiply if other types of eventual distributions are assumed and the evidence for them must be obtained from the original series. In most instances, some information relating to distribution type can be obtained on theoretical grounds. In order to be in a position to apply theory developed for the simple or normal law type, interest centers in normalizing transformations. Tukey[11] recently made an important contribution to this subject. There is a need, however, for papers relating to adaptation of transformations for use in Quality Control.

SUMMARY

This paper reviews in a general way the use of statistics and quality control in presenting historical data, in making discoveries and in making predictions. Certain weaknesses in many predictions are traced to failure to make tests relating to the assumptions on which these predictions are based. Illustrations are given of the use of such tests in a few simple cases and the modification in the predictions resulting therefrom. Finally, a way for providing approximate probabilistic measures for such treatment is outlined. Thus, statistical evaluation is justified on engineering grounds where the assumption of statistical control is untenable unless it can be demonstrated.

REFERENCES

(1) Membership Directory 1958, P. XIX, American Statistical Association, Washington, D.C.

(2) The Contributions of Karl Pearson, Helen M. Walker, Journal of the American Statistical Association, 53, 11-22, March 1958.

(3) Economic Control of Quality of Manufactured Product, W. A. Shewhart, D. Van Nostrand Company, New York, 1931.

(4) On the Mathematical Foundations of Statistics, R. A. Fisher, Philosophical Transactions, London, A222, 309-68, April 1922.

(5) Statistical Method from the Viewpoint of Quality Control, W. A. Shewhart, The Graduate School, Department of Agriculture, Washington, 1939.

(6) How to Detect the Type of an Assignable Cause, P. S. Olmstead, Industrial Quality Control, 9, 32-4, 36, November 1952 and 9, 22, 26, 28, 32, January 1953.

(7) Control Charts Using Midranges and Medians, E. B. Ferrell, Industrial Quality Control, 9, 30-32, March 1953.

(8) Runs Determined in a Sample by an Arbitrary Cut, P. S. Olmstead, Bell System Technical Journal, 37, 55-82, January 1958.

(9) On the Analysis of Planned Experiments, M. E. Terry, American Society for Quality Control, National Convention Transactions, 553-6, 1955.

(10) Properties of Control Chart Zone Tests, S. W. Roberts, Bell System Technical Journal 37, 83-114, January 1958.

(11) On the Comparative Anatomy of Transformations, J. W. Tukey, Annals of Mathematical Statistics, 28, 602-32, September 1957.

(12) The relation Between Confidence Intervals and Tests of Significance, M. G. Natrella, The American Statistician, 20-22, February 1960.

A PROGRAM FOR DEFINING AND CONTROLLING VISUAL DEFECTS IN A MULTI-PLANT OPERATION

E. E. Allen
Manager, Quality Control Division
American Can Company

When I attended my first basic course in Statistical Quality Control about 10 years ago, I was told of the great advantages of this new technique in controlling the quality of products which were made at high speed in large volume. This appeared to be right down my alley as cans were produced at the rate of 20,000 to 24,000 per hour. However, from that point on until this day, I have seldom heard about high speed in connection with statistical methods. Apparently, this type of application was so simple as to not require explanation or example. At this course, also, it was stated that considerable savings were possible through S.Q.C. by the elimination of 100% inspection but we had very little of such inspection. We were instructed that variables control required a sample of 2 to 5 parts from a process at each inspection but with as many as 8 or 9 separate duplicate processes in a single production line, we were taking only a one can sample from each. Did these methods require that we increase our inspection cost? Besides, we were interested in within part as well as part to part variation in many cases. How were we to statistically control individual measurements to stay within specifications? Even the experts were at a bit of a loss to give us practical advice on these problems. Well, suffice it to say that after working with S.Q.C. for a time, we came to believe in it and have since made sufficient progress to amply justify this belief.

I am here today to tell you about one phase of that progress -- the visual inspection of cans and can parts. I hope that some of the things we have done to improve our inspection will be of interest and perhaps assistance to you in solving some of your problems.

Background

Prior to the advent of S.Q.C. in the American Can Company, visual inspection was a patrol type procedure done by inspectors who visited each line once per hour, looked at 6 to 10 cans and decided whether the quality being produced was satisfactory, was in need of corrective action or was so bad that the line should be shut down (by the inspector) and a rejection made. The line operating people, of course, did some visual inspection and there was an automatic tester in the line for rejection of leakers. This inspection had been sufficient to avoid most serious complaints from customers because at that time they were primarily concerned about cans that leaked or would not go through their filling and closing operations.

Customers' attitudes, however, were rapidly changing. More internally enamel lined and more externally lithographed containers were coming into the picture; new products were being developed which made greater demands on the cans; and users of all products were becoming much more critical of quality. Where formerly only rather high percentages of major type defects resulted in rejections by customers, now low percentages of both major and minor defects were viewed with real concern. Therefore, it was evident that our inspection would have to be overhauled to provide more accurate information at lower defect levels. Statistical methods offered an answer to this need.

Defect Standards

Before any visual inspection method can be really effective in providing uniform estimates of quality, it is necessary that there be a uniform basis for grading individual items inspected as acceptable or defective. Without uniform standards, percentages of defects found by inspectors would have no meaning. Cans

were being graded in the American Can Company but there were fairly wide differences of opinion in many cases as to how bad a particular irregularity had to be to be called a defect and often circumstances such as production quality problems or meeting of shipping schedules, further affected this judgment. Therefore, it was obvious that defect standards must be developed.

In arriving at standards, there were four questions which required answering.

1. Where is the line to be drawn for each type of defect?

2. How are the defects to be classified relative to seriousness?

3. What consideration is to be given to irregularities not severe enough to be called defects?

4. How can the standards established be made available to everyone requiring them?

To establish standards for each type of defect, it was decided that where severity of irregularity would be the deciding factor, the standard would represent the point at which the irregularity was just severe enough to grade the container defective. It was recognized, of course, that in the case of certain types of defects, if the irregularity was present at all, the container would be considered defective.

A list of the types of defects encountered were then developed and selected plants were asked to gather samples of cans and covers illustrating all types of defects at various levels of severity. These samples were sent to a central location for consideration and selection of individual defect standards. A meeting was then set up with key people from the Quality Control Division (knowledge of customer preferences), the Technical Service Division (knowledge of performance of containers in the field) and the Sales & Service Division (knowledge of requirements for use in customers' equipment). At this meeting, each type of defect was considered separately and a defect standard sample selected from those illustrating the various levels of severity. Then a defect definition defining the sample was agreed upon. In some cases, satisfactory defect standard samples could not be selected from the various samples available, so defect definitions were agreed on and a sample to fit the definition obtained later. There were instances also, where it was not possible to define a defect in words and in those cases the standard sample had to define as well as illustrate the defect. Although this procedure for setting defect standards may appear simple, a considerable amount of work was involved before a satisfactory set of defect standard samples was finally obtained and much discussion was needed to obtain agreement on both samples and definitions.

With regard to classification of defect standards, it was recognized that some defects are more serious in effect than others and consequently, should be separated from the others so that they might be treated independently. Therefore, three classifications -- Major, Minor and Marginal, were established and defined.

Major - Defects which are certain or likely to cause failure of the can to fill its intended purpose, i.e. fail to hold the product, to perform satisfactorily in the customer's equipment or to be saleable to the ultimate consumer.

Minor - Defects which are definitely objectionable but will not cause failure of the can to fulfill its intended purpose.

Marginal - All irregularities less severe than the visual defect standards.

Each type of visual defect was considered individually by the same group that agreed on the defect standards and placed in the Major or Minor classification. Marginal irregularities were automatically classified by definition.

The marginal classification was included for two reasons, (1) these irregularities can develop into actual defects, if allowed to go uncorrected, and (2) if a high percentage of cans have such irregularities, they could become objectionable to the customer. Consequently, it was decided that a count of them should be maintained so that corrective action could be taken should they exceed a set limit.

Although this completed the establishment of visual defect standards, the problem still remained as to how these standards could be made available to those who were to use them. With a large number of plants and other units requiring them, it was found to be next to impossible to select enough duplicate sets of containers and covers for illustrating the defects. Anyway, such samples would not be permanent as they would deteriorate over a period of time.

Regular photographs were considered but rejected as color and depth were required in illustrating certain defects. The Army Quartermaster Corps had used special photographs to show can dent standards which produced a three-dimensional image when viewed through polaroid glasses. This method had two disadvantages, however, they could not be made in color and they tended to deteriorate in time. This led us to investigate 3-D stereo slides in color which proved to be very satisfactory. All defect standards were photographed by this method and set in 3-D slides. Strip-ins were included at the bottom of each photograph giving (1) the type of defect, (2) the defect definition, (3) the word "limit" unless just presence of the irregularity in any severity determined defectiveness and (4) the date the slide was prepared so that revisions could be identified (see Figure 1). Each slide was also identified with a label between the two pictures. Besides the type of defect, the label included a code identifying the type of container, the category of defect (L for lithograph, S for solder, etc.) and the number of the defect in the category.

Inspection and Control

With defect standards established, the next step was to set up standard procedures covering methods for inspection, charting and feed-back. It was desirable that the methods used be as simple and uncomplicated as possible to maintain; that data gathered be shown on charts in a manner easy to interpret; and that the most information be derived for the dollars spent. The major requirement which had been placed on the entire S.Q.C. program by Management was that it not increase the cost of inspection.

After preliminary development work over a period, at a number of plants using various sampling routines and several modifications of p and np charts and after making some changes in responsibility for in-process inspection, a standard procedure was set up. This standard procedure had to be written, however, to allow flexibility to handle varying conditions which occur from plant to plant so requirements affecting amount of inspection were placed on a minimum required basis with instructions that they should be increased if necessary for control or protection. As it was necessary that inspectors handle more than one line and also, that certain items be checked which required destruction of the can, completely random samples could not be obtained. Therefore, a minimum of 4 sub-groups were required to be taken at intervals throughout the sample period alternating between lines on the completion of each sub-group. The destructive checks were to be made on at least one can from each sub-group. The minimum sample size was set at 100 and the lot size was not to exceed one or two hour's production depending on the type of container being run.

The charting method for data developed by this inspection was to be recorded by defect at the bottom of a special form developed for the purpose (see Chart I). Minor defect totals only were to be plotted on a p chart at the top of the form. Rejection limits were to be drawn in on the chart for minors and written

in for each type of major defect to the right of the defect name at the bottom of the form. Corrective action limits were to be written in for marginal defects at the bottom of the form. A rejection was to be made and corrective action requested when the rejection limit on the chart for minors or the rejection limits for any individual major defect was exceeded. Corrective action only was to be requested when any individual marginal irregularity exceeded its limit. As marginal irregularities are in most every case less severe irregularities of the same types as major and minor defects, a circle was to be entered in the proper space opposite the type of defect in the lists of these defects to indicate which marginal irregularity required action when the limit was exceeded. All reject action is noted on the chart to identify the run with the rejected cans. Periodically, defect data is summarized by types of defect to bring continuing trouble spots to the attention of Management.

Administration

The establishment of a usable set of defect standards, together with efficient and workable procedures for sampling, recording and analyzing inspection data, although important to a visual inspection program, will not guarantee its effectiveness. It must, also be properly administered or it will fail.

There are three main areas where error can cause trouble in the operation of a visual sampling plan (1) grading of samples, (2) recording of data (including computations) and (3) making decisions. Of these, the most prevalent error occurs in the grading of samples.

We have found that inaccurate grading is generally the result of one or more of these factors.

a. Honest errors in the interpretation of defect standards. This can be caused by inherent poor judgment but is usually because of poor training.

b. Tendency to relax a standard defect definition. This is caused by trouble being experienced with a type of defect over a period with the result that percentages would be continually high. This can happen only with some of the minor defects if Management allows a line to run for a time before correction.

c. Carelessness. This is the result of an inherent characteristic of the inspector or a lack of proper disciplinary action.

d. Brainwashing by operating people charged with taking required corrective action. This is usually the result of a clash between a strong personality and one that is weaker.

The problem, however, is not so much in finding out why an inspector is grading inaccurately but in how inaccurate practices can be caught and corrected before a customer rejection is encountered.

As a part of our visual inspection program, we have required that the Quality Control Department in each plant maintain a program for auditing inspectors. Inspectors are required to hold each sub-group until they have finished with it and at a minimum of once per week -- more frequently if the inspector is new or known to need improvement -- a member of the supervisory staff is to re-inspect one of these sub-groups from each inspector. The re-inspection results check the inspector directly. All re-inspections are recorded and summarized periodically to compare with the long-term average percent defective found by the inspector. Comparisons between shifts on the same lines, compare inspectors on these shifts and are used to indicate sources of inaccuracies. Sampling of product going to the warehouse may be used to check inspectors if carelessness is suspected.

As differences in interpretation of defect standards between plants can occur and may affect the visual quality of cans being shipped to common customers, the Quality Control Division in the General Office conducts a plant audit on a yearly basis. Two or more lines are checked by re-inspection of all samples checked by inspectors assigned to these lines on each of two shifts. As the inspectors know the audit is being made, differences are assumed to be honest differences in interpretation. If previous results on all of the same lines with the same inspectors are significantly different from those found, it is suspected that the inspectors are not doing their best job and the plant Quality Control Supervisor is expected to investigate and take any necessary action to improve accuracy.

From a company standpoint, the results of these audits are of considerable value as a check on the adequacy of the visual standards. Also, based on information gathered in these audits, a number of standards have been revised and expanded to improve ease of interpretation.

Summary

Effective inspection and control of visual defects requires:-

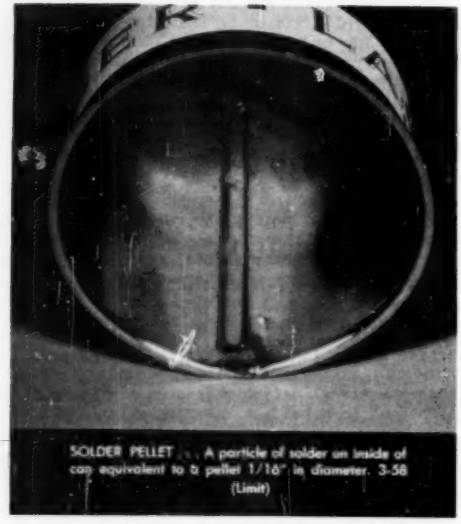
1. Adequate visual defect standards available at all work locations in a readily interpretable form.

2. Efficient and easily understood inspection, recording and analysis procedures written up and available at locations where they are to be used.

3. Well laid out and effective inspector training procedures.

4. Good supervision plus an inspection auditing program to periodically check whether inspection is accurately and uniformly performing its function.

5. Prompt corrective action on quality deviations when they occur.



SOLDER PELLET . . . A particle of solder on inside of can equivalent to a pellet 1/16" in diameter. 3-58 (Limit)

COMPOUND SKIP . . . Any excessive narrowing of the film on the flat of the flange. 3-58 (Limit)

FIGURE 1

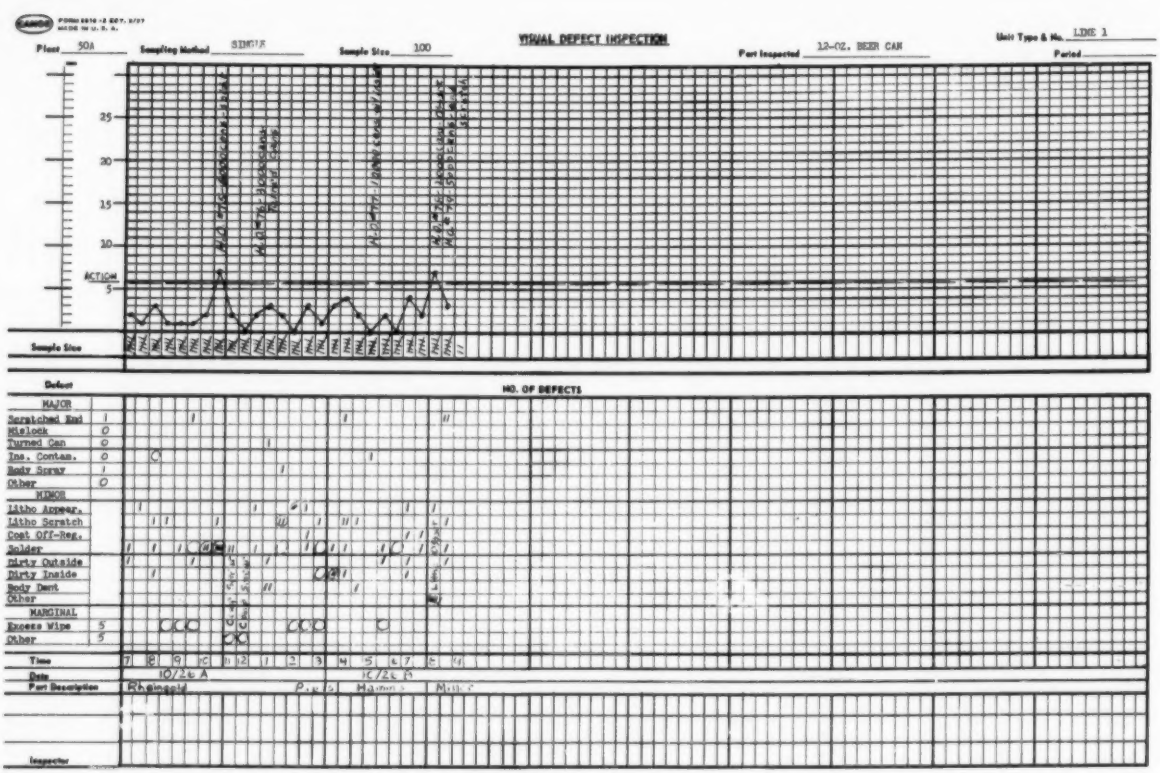

VISUAL DEFECT INSPECTION
Plant 50A
Sampling Method SINGLE
Sample Size 100
Part Inspected 12-OZ. BEER CAN
Unit Type & No. LINE 1
Period
25
20
15
10
ACTION
5
Sample Size
Defect
NO. OF DEFECTS
MAJOR
Scratched End 1
Mislock 0
Turned Can 0
Ins. Contam. 0
Body Spray 1
Other 0
MINOR
Litho Appear.
Litho Scratch
Coat Off-Reg.
Solder
Dirty Outside
Dirty Inside
Body Dent
Other
MARGINAL
Excess Wipe 5
Other 5
Time
Date
10/24 A
10/25 A
Part Description
Rhinegold
Inspector

CHART I

RELIABILITY IN COMMUNICATIONS AND ADMINISTRATIVE APPLICATIONS

Harry G. Romig, Reliability Director, Engineering
Military Products Division, Hoffman Electronics Corporation

NATURE OF PROBLEM

Management must carry out its functions effectively and be as efficient as the line organizations. The Board of Directors establishes policies which must be carried out by the administrative groups. Some will be covered by Executive Directives while others will be long-range programs that must be developed and are subject to change as the program advances with time. It is necessary therefore to secure the best and most reliable lines of communications and to have an organization that reacts rapidly to the orders as received.

Ordinarily Managements believe that a few simple directives will be all that are necessary in order to have the basic organizations function as desired. In the past a few errors have occurred but little effort was made to determine the cause. Now interest has been aroused in Human Engineering and steps are being taken to analyse these mistakes and see what measures should be taken to avoid them. In order to understand the nature of the problem and all the various items that must be taken into consideration, specific cases must be studied. This paper will take up for consideration many types of problems and describe how better communications and improved organizations might decrease the mistakes and increase the efficiency of the operations.

The fields of activities must be specified. Educational institutions have different basic problems than Insurance Companies and similar industrial organizations. Stores, banks, service plants, those engaged in military production, commercial activities, newspapers, radio and television broadcasting stations, manufacturing divisions, mail order shops should be included in the fields of activities as well as distribution problems involving domestic and foreign trade or possibly job shops, strategies in business competition, forecasts of business conditions and the needs of particular industries and plants. Particular attention should be paid to administrative problems and means of communication. Each field has its own problems but there are many similar problems that may be handled as a group with a common solution. Consideration might even be given to the efficient operation of technical societies and how to maintain the reliability of the communications in such societies. It is apparent from these few examples that there are many varied problems and activities. However, there does exist some pattern of behavior in this complex society and these various fields, which should make it possible to provide some common solution that will provide for reliable communications with minimum errors.

RELIABILITY TECHNIQUES

Having outlined the nature of the problems that are to be considered, it next must be noted that techniques will be applied to solve some of these various problems. Our National Quality Control organization prides itself on the fact that it has basic knowledge concerning the Applications of Statistical Quality Control to production and allied areas. It appeared for a time that Quality Control was to be used synonymous with Inspection, particularly in manufacturing plants and especially in those dealing with government contracts. The Military Divisions took their Inspection Departments and renamed them Quality Control Departments. As a result very little work in true Quality Engineering was done and naturally applications to Administrative problems occurred only in rare instances.

The first work in this area by Dr. Walter A. Shewhart, the father of Quality Control, dealt with the Control of Quality of Manufactured Products. His Washington lectures, edited by Dr. W. Edwards Deming, gave the basic philosophy of Quality Control and expanded the field of applications to all the areas and even more fields than those enumerated above in the discussion of the fields

of activities involved. Since then Quality Control has meant many things to many different people. It has tried with little success to be recognized as a truly professional activity, the same as law and medicine. Recently many have tried to again throttle it within a small domain. The emphasis on Reliability by the Military and edicts that Reliability shall have the last word on all items shipped makes Quality Control subservient to Reliability. Possibly this is the best solution. To counteract these trends, Dr. A. V. Feigenbaum has prepared for publication and given lectures on Total Quality Control as he sees it. However, he makes no mention of Dr. Shewhart's work in this area which includes all of Total Quality Control and more areas in Research and Development. Also a Committee on Professionalism has worked out a set of principles to assist in gaining recognition in the various States of Quality Control Engineers as truly Professional Engineers. A Professional Engineers License for Quality Control is desired by a few. Some Universities are giving more and more courses in Quality Control and allied subjects so that its wider applications may be covered and these doctrines may be spread into many divergent areas.

Two other types of controls have been added with some variants. Industrial Engineering in some instances has added all the tasks of Quality Control to its areas of activities so that Quality Control in many companies may be a part of Industrial Engineering. This has been very successful. Operations Research has taken the long-range view of management problems and covers not only all the Quality Control problems but any type of problem that Management must handle. These two disciplines, Industrial Engineering and Operations Research are given recognition by the Universities and degrees in these subjects are given. Since there are no degrees given in Quality Control, one may see readily why Managements are willing to establish at a high level operations in Industrial Engineering in some cases, and also in Operations Research in the larger companies.

The casual reader probably expected this exposition to name as basic disciplines Reliability and Quality Assurance. These are variants. Even though Directors of Reliability and Quality Assurance exist in many companies, there is no reason why these may not be merged with Quality Control, Inspection, and Test and provide a professional discipline known as Quality Engineering or Control Engineering. This Quality Engineering would contain all the elements needed for the successful control of all phases of the operations required not only to develope, produce and maintain quality items, but also to control the administration of the organizations that achieve these goals.

It is realized that Quality Control broke down with respect to the attainment by the Military of Reliable Products. Maintenance costs have been found so high that steps are being taken to provide some Figure of Merit. One study claimed that some units required 10 Spares per year to keep in operation. One need only see the orders for spares, modification kits, rework, and compare these with the original orders to see that reliable units are few and far between. Hardware, i.e. Equipment Units, Modules, and Assemblies, is made, but a large part have a high probability of failure value. Reliability has been demanded as a contract requirement. Reliability in the Military sense is associated with the successful conclusion of a mission. Now this mission may be short, taking only a few seconds or a fraction of a second, or the mission may be years, such as is planned for these solar batteries on Pioneer V, from which messages are expected to be obtained more than three years from now. The AGREE Report provides for tests at different stress levels to check the reliability of certain systems and assemblies, and the components used in their construction. Probabilities of failure, constant failure rates, evaluations of Mean Time Between Failures (MTBF) are some of the techniques used for Research and Development as well as for Production. Associated with these Reliability values are Probabilities that measure the Confidence Limits, or more aptly we might say the Confidence Bandsassociated with given Probability values. What Assurance or Degree of Belief do we have that satisfactory performance has been achieved? This question and many others must be answered in order that our application of Reliability techniques may be correct and be recognized as good standard practices.

Critical experiments should be designed, parts and components must be studied and analysed to determine their contribution from a reliability point of view to the overall picture. Rating systems must be devised that will make it possible to evaluate this work in a reasonable fashion and reduce costs by emphasizing the use of these qualified nonstandard parts that apparently function as desired. For paper-work evaluations factors in Human Engineering should be studied so that account is taken of the fatigue factor. Some have found that 95% of the troubles experienced were due to human errors of some kind and/or type. Audit inspections may be taken and the results posted. However, care must be exercised. Reliability techniques include a better procedure for handling problems and establishing controls that are truly effective.

COMMUNICATIONS FLOW CHARTS

Communication systems are quite similar to the nervous systems of mankind and even the lowest forms of animal life. These are outlined rather completely in the paper. Examples are given to show these in detail and also to indicate how to make such systems reliable by the application of the most recent Quality Control and Reliability techniques.

WEEKLY OR DAILY BOARD MEETINGS — EXECUTIVE COMMITTEE CONTROLS

Considering the typical business, the paper discusses the type of information that Executives desire for their daily, weekly, and also monthly and yearly analysis of current operations. Reliable communication techniques must be employed in these areas. The paper illustrates the various problems by case studies. The paper details such Board meetings or their counter-part, the Executive meetings. The paper gives examples in these areas and indicates what should be done under ideal conditions. These examples may then be compared with actual life data. Stream-lined information and data will be discussed and its opposite, the usual inefficient communication center and system.

SCIENTIFIC HANDLING OF MANAGEMENT PROBLEMS

Management often has difficulty with the various prototypes during design and fabrication, but also with the production models. An Operations Research approach proves useful and is discussed in the paper.

EFFECTIVE ORGANIZATIONS

Organizations vary in their complexity. In small companies the organization best suited is simple and very flexible. In larger concerns, staff functions become more important. Other parts, such as those from stock with and without inspection, must be evaluated. Current procedures are rather lax. Efficient organizations are obtained by good hard, rational work. Examples as well as this theory are developed fully in the paper.

RELIABILITY IN ADMINISTRATIVE ACTIVITIES

Areas requiring consideration are noted. Techniques of Operations Research and Reliability are applied. Case studies indicate the efficiency of the various systems and that the Reliability approach has considerable merit. This Section includes the practical applications of these statistical and remedial measures. The results are summarized. References are also given where needed.

50,000 PARTNERS

C. W. Harris, Director, Quality Assurance & Test Services
Lockheed Missiles & Space Division
Sunnyvale, California

It is indeed a pleasure to speak before the ASQC at their 14th Annual Convention, and particularly so since it is the first such conference held west of the Mississippi River. Although it has been a long time since the statement "Go West Young Man, Go West" was first made by Horace Greeley, it just shows that sooner or later young men do go west. The subject of statements brings me to the introduction of my paper, which deals with standard quality assurance and reliability statements. I venture to say that not one of us has attended an ASQC meeting or any reliability symposium in which we did not hear the following: "We survey all of our suppliers before a purchase order is placed", or "Our Quality Assurance organization only approves purchase orders from suppliers which have previously been approved for adequate reliability and quality assurance programs", or "Our company supplies our vendors with test equipment specifications and test procedures with the purchase order", or "All of our purchase orders now contain specific reliability requirements", or "All our vendors are required to demonstrate that the calibration of their test instruments tie into the National Bureau of Standards Primary Standards Laboratory in their area".

I would like to relate how difficult it is to live up to these statements on the research, development, and initial production of two fast moving DX priority weapon system programs. To place the difficulty of complete compliance under these priority conditions in the proper light and at the risk of possibly permitting a direct comparison between our organization with those companies who thoroughly apply the quality assurance axioms, I will show the magnitude of the problem through the presentation of pertinent statistics taken from the records of the Lockheed Missiles and Space Division.

Last year, we did approximately 500 million dollars worth of business, however, 250 million, or half of this amount, was paid to our subcontractors and suppliers. Of this 250 million we have approximately 57 large subcontractors. Disregarding for the moment the status of their quality assurance and reliability programs, we find that approximately 150 million dollars was paid to what is generally considered small businesses, jobbers, warehouses, and supply agencies. The latter amount of business was placed through the media of approximately 125,000 purchase orders. Approximately 5% of these purchase orders called for LMSD source inspection. Now, to further bring out the immensity of the problem of strict adherence to the commandments mentioned above, our statistics show that these purchase orders involved 6,000 different suppliers. Further, each one of the organizations in which we place a purchase order has an average of 8 to 10 sub-tier suppliers. By simple arithmetic it then shows that in order to make good some of the statements we are making at these meetings with respect to supplier surveys, supplier quality assurance and reliability program approval, provision of test equipment specifications and procedures, specific reliability requirements, and documented calibration, we would have to be sure that 50,000 organizations are considered. To further draw on our experience at LMSD, it is significant to note that this type of surveillance would have to be done on a very tight R&D and production schedule. For example, the Polaris program was, from its inception in the summer of 1956, to reach Initial Operational Capability in 7 years, i.e., summer of 1963. However, as you know, the Initial Operational Capability of this program is scheduled for the fall of 1960. A schedule compression of approximately 34 months or in other words 42% less time than we originally contemplated. The tightening of the schedule on our Satellite Systems effort, our other major program, is of the same magnitude. In view of the volume of work that is subcontracted and the ever tightening schedule requirements for its completion, it is redundant to say that many of the quality

assurance commandments are based upon desired operating policy, rather than a set of hard and fast rules by which we operate.

Now, since our total organization is dedicated to meeting scheduled production of only those weapon systems which have the highest level of reliability and quality, we can immediately see that the Quality Assurance organization is in a very difficult position. Furthermore, it is reasonable to assume that the tightening of the schedule will immediately reveal late design releases. To accentuate the quality assurance position, let us examine for a moment how the purchasing organization reacts to these conditions. First of all, only a limited number of supplier organizations have the capability, equipment and talent to perform some of the research and development tasks. This condition invariably leads to the selection of sole sources. Further, past experience shows that during the early phases of the design, due to the R&D nature of the program and changes brought about by repeated schedule accelerations, that the reliability experts are unable to fix on a reasonable and acceptable quantative reliability factor. Where reliability factors can become a part of the supplier's control, the prime contractor is hard pressed to specify the manner in which they will require the supplier to demonstrate compliance. Admittedly, the engineering organization could be safe and insist that the supplier have 100% or nearly so reliability, but this restricts the suppliers to a point where they can be counted on one hand. Let's look further now into the frantic life of the purchasing agent. We find that placing the purchase order through even the minimum routine channels is too time consuming so he resorts to telephoning the purchase order, promising of course on bended knee that the paperwork will follow immediately. I am sure that all of us engaged in quality assurance know how fast immediately is in these instances. In fact there have been some cases where parts had arrived on the receiving dock and were turned over to quality assurance before the purchase orders were cut. To complete the picture, let me mention the confusion wrought through the dubious assistance sometimes given the purchasing department by the engineers' acceptance of the suppliers' hand-delivered components and subassemblies.

I would like to say a little bit about the difficulties of applying these oft-repeated statements regarding supplier quality assurance and reliability programs to the spares problem. Unfortunately, it seems that the spares for the suppliers' sub-assemblies are furnished by one or more of the 8-10 subs mentioned earlier. It is not uncommon in the press of the program for the engineer to order belts, components, or blackboxes for a suppliers' assembly via the only identification available, namely a photograph or better yet, unit #3a shown in the left hand corner of page 6 in the supplier's catalog. Obviously, proper inspection or certification of these parts on a timely basis can not follow the usual procedure.

For those who recognize that these problems to one degree or another exist in their own organizations, it is evident that we can not always wait until all the data is available to the supplier before we place the purchase order, or can we control the host of suppliers by sending hundreds and hundreds of reliability/quality assurance personnel out into the field. Therefore, we must recognize that the prime contractor is relying heavily on the integrity and sense of responsibility of the subcontractors and suppliers to achieve the highest product reliability within the cost and schedule established by the purchase orders.

This situation makes the supplier and his subs, 50,000 of them, partners with LMSD in the management of the weapon system. Having accepted them as partners, we must help them meet their schedules with quality products. We must also bear in mind that some of our partners do not have a statistical quality control function or ever heard of the classic theories of probability or can guarantee with a confidence level of 90% that the reliability of their product will be 95%. In recognition of the typical suppliers' problems on one hand, and the R&D nature of the programs, as well as the tight schedule on the other hand, we have placed considerable emphasis and effort on thoroughly evaluating the "as received" state of their products. Where product quality is good we are delegating some Material Review Board action and control of material changes to the supplier's organization. To assist him in achieving maximum reliability, we are

supporting him where necessary with our own test laboratory services, reliability data, and technical assistance. An extensive qualification and finally requalification program on parts selected at random from each lot shipment is the cornerstone of our evaluation process. This method though costly has significant advantages over attempted field control part and sub-assembly surveillance. It provides "on the spot" engineering analysis of the "as received" assembly and permits the application of the latest environmental stress information, which may have just arrived from the test bases. Functional failures of randomly selected samples in the requalification test many times detect changes in processing and part installation by the supplier which defy detection by field personnel at the supplier's plant. In addition, the possibility of rejecting an entire lot of a supplier's completed component, assembly, or system has far more impact on his organization than correspondence urging him to improve his documentation or processing methods.

It is realized that the advantage of evaluating the supplier's product in the "as received" condition carries with it a penalty of determining the discrepancy at a later date than we would like to accept. To prevent the receipt of an uncontrolled number of unsatisfactory units, we have activated our Quality Assurance field operations. This organization, with a Manager at Sunnyvale, operates four regional offices, which are located at Sunnyvale, Van Nuys, Chicago and New York. The major effort of personnel permanently assigned to these offices is expended in performing the required source inspection and selected surveys, as well as providing media for "on the spot" assurance that corrective action is being taken by the supplier, as a result of discrepancies found by the in-plant evaluation program. Where the volume of work or past performance indicates marginal product quality, permanent Quality Assurance representatives are assigned to a given supplier. Here again, it is obvious that the assignment of a single individual or two to a large supplier will not in itself assure that all processes, inspection, and test results meet the purchase order requirements. We recognize that such an assignment will accomplish more if that organization's management is apprised of and assumes the responsibility for a sound reliability/quality assurance program, providing the proper documentation and corrective action systems. The degree to which the Quality Assurance representative accomplishes these objectives can best be measured by the improvement of the supplier's quality level, which in any organization can only be achieved through general management's support and interest. Another function of the field office which should be examined is that performed by the itinerant Quality Assurance source inspection representative, whose effectiveness hinges upon two probabilities. One is the probability that he arrives at the supplier's plant when some deviation to the quality assurance requirement is occurring and the other is that he will note this deviation. Again, it follows that the integrity and sense of responsibility of the supplier is his most important attribute. We consider that the company representative's success in impressing the supplier's management with these facts is equal in importance to affixing his signature to the "release to ship" paper. Since this type of supplier surveillance draws upon the general background and experience of the Quality Assurance representative, rather than his detailed knowledge of the particular component, we have found it desirable to re-cycle these personnel through the Sunnyvale facilities at least twice a year for indoctrination and reorientation by the major design organizations. Where quality problems indicate that the supplier is having basic design or manufacturing difficulties, the Quality Assurance representative requests assistance from the appropriate Sunnyvale organization. This method of operation has proven to be the fastest and most economical way of helping the supplier solve a particular problem and places the field Quality Assurance representative in the proper relation to the designer at Sunnyvale and the supplier's personnel.

The supplier problems described in the preceding pages are typical of those encountered by Quality Assurance personnel during the research, development, and initial production of tightly scheduled DX priority weapon systems. During this phase of weapon system management, the solution to these problems by the application of conventional controls and procedures would, in our own experience, have been unacceptable to our customers from the standpoint of available money and time. Accordingly we believe that thoroughly checking the partner's product

in the "as received" condition, where the proper test equipment and latest test procedures are available, yields the most effective evaluation of the unit and provides a specific basis for the supplier's immediate corrective action. We also believe that the best results from the Quality Assurance field operations can be obtained initially by emphasizing enforcement of corrective action, rather than large scale surveys and repeated source inspection visits.

It is recognized that if we are to achieve full production, we must place follow-on orders with suppliers, who have proven satisfactory during the R&D and initial production evaluation. At this point, we are able to include in the conditions of the purchase order the normal reliability factors, quality control procedures and acceptance demonstration requirements. However, we can not get away from the fact that regardless of the purchase order conditions, the successful performance of our 50,000 partners is still primarily dependent upon their continued integrity and sense of responsibility for the production of a reliable product.